

Jarrell, Noble

182151

From: Ramirez, Delia
Sent: Tuesday, March 14, 2006 1:44 PM
To: Jarrell, Noble
Subject: 10/673786

Hi,

I would like to request the following search:

1. SEQ ID NO:1-2 in the nucleic acid databases (commercial and interference)
2. SEQ ID NO:2 in the protein databases (commercial and interference)

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

Noble
SPR Fin 3/17/06
SONL
2NA Compaugnan
1AA

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 14:55:19 ; Search time 2975 Seconds
(without alignments)
3310.531 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191

Sequence: 1 atgtttgagaacattaccgc.....cgattgtgcagtgctgttaa 1191

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1191	100.0	1191	6	US-10-369-493-24472	Sequence 24472, A
2	1191	100.0	1191	7	US-10-673-786A-1	Sequence 1, Appli
3	1191	100.0	14759	9	US-10-893-671-87	Sequence 87, Appl
4	655.8	55.1	1188	6	US-10-369-493-44812	Sequence 44812, A
5	510	42.8	966	6	US-10-369-493-23918	Sequence 23918, A
6	496.2	41.7	2133	9	US-10-450-763-23995	Sequence 23995, A
7	486.2	41.7	3222	9	US-10-450-763-25886	Sequence 25886, A
8	493.4	41.4	7977	9	US-10-795-159-536	Sequence 536, App
9	493.4	41.4	908766	9	US-10-795-159-685	Sequence 685, App
10	471.8	39.6	1191	7	US-10-275-026A-187	Sequence 187, App
11	466.8	39.2	17381	9	US-10-915-740A-40	Sequence 40, Appl
12	466.8	39.2	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
13	458.2	38.5	1830121	7	US-10-329-670-1	Sequence 1, Appli
14	458.2	38.5	1830121	8	US-10-158-865-1	Sequence 1, Appli
15	458.2	38.5	1830121	9	US-10-981-687-1	Sequence 1, Appli
16	363	30.5	2220	6	US-10-450-763-17608	Sequence 17608, A
17	331.8	27.9	1185	6	US-10-369-493-35511	Sequence 35511, A
18	327.8	27.5	1191	6	US-10-369-493-32111	Sequence 32111, A
19	320.8	26.9	1188	6	US-10-369-493-31184	Sequence 31184, A
20	320.8	26.9	1194	6	US-10-369-493-28425	Sequence 28425, A
21	310.6	26.1	1194	6	US-10-369-493-37712	Sequence 37712, A
22	306.6	25.7	24417	6	US-10-216-209-1	Sequence 1, Appli
23	297.4	25.0	1093	9	US-10-450-763-13154	Sequence 13154, A

ALIGNMENTS

RESULT 1

US-10-369-493-24472
; Sequence 24472, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24472

Query Match	100.0%;	Score 1191;	DB 6;	Length 1191;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1191;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGTTTGAGAACATTACCGCGCTCTCTGCGACCCCGATTCTGGGCCCTGGCCGATCTGTTT	60	
Db	1	ATGTTTGAGAACATTACCGCGCTCTCTGCGACCCCGATTCTGGGCCCTGGCCGATCTGTTT	60	
QY	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG	120	
Db	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG	120	
QY	121	GGCAAAACCCGGTACTGACCGCTGAAACAGGCTGCAACAGTATCTGCTCGAAATGAA	180	
Db	121	GGCAAAACCCGGTACTGACCGCTGAAACAGGCTGCAACAGTATCTGCTCGAAATGAA	180	
QY	181	ACCACCAAAATTAACCTCGGCATTGACGGCATCCCTGAAATTTGGTGCCTGCACCTCAGGAA	240	
Db	181	ACCACCAAAATTAACCTCGGCATTGACGGCATCCCTGAAATTTGGTGCCTGCACCTCAGGAA	240	
QY	241	CTGCTGTTTTGGTAAAGGTAGCCCTTGATCAATGACAAACGTCGCTCGCACGGCACAGACT	300	

```

241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTGCACGGCACAGACT 300
301 CCGGGGGGCACTGGCGCACTACGGCTGGCTGCGGATTTCTTGGCAAAAAATACAGCGTT 360
301 CCGGGGGGCACTGGCGCACTACGGCTGGCTGCGGATTTCTTGGCAAAAAATACAGCGTT 360
361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATCGGAAATACACTCTTGAATTCGAT 480
421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATCGGAAATACACTCTTGAATTCGAT 480
481 GCATCTGATTTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 540
481 GCATCTGATTTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 540
541 TGCCATTAACCAACCCGCTATCGCCCTACGCTGGAACAATGGCAACACTGGCACAACTC 600
541 TGCCATTAACCAACCCGCTATCGCCCTACGCTGGAACAATGGCAACACTGGCACAACTC 600
601 TCCGTTGAGAAAGGCTGTGTTACCGCTGTTGACTTTCGCTTACCAGGGTTTTGGCCGCGGT 660
601 TCCGTTGAGAAAGGCTGTGTTACCGCTGTTGACTTTCGCTTACCAGGGTTTTGGCCGCGGT 660
661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGCTATGCAATAAAGAGCTGATTTGTT 720
661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGCTATGCAATAAAGAGCTGATTTGTT 720
721 GCCAGTTCCTACTCTAAACCTTTGGGCTGTGACAGAGCGTGTGGCGCTTGTACTCTG 780
721 GCCAGTTCCTACTCTAAACCTTTGGGCTGTGACAGAGCGTGTGGCGCTTGTACTCTG 780
781 GTTCTGCGCAGTGAACCGTTGATCGCGCAATTCAGCCAAATGAAAGCGGCGATTCGC 840
781 GTTCTGCGCAGTGAACCGTTGATCGCGCAATTCAGCCAAATGAAAGCGGCGATTCGC 840
841 GCTAACTACTTAACCCACCAAGCAACGCGGCTTCTGTTGTCACCACTCTGAGCAAC 900
841 GCTAACTACTTAACCCACCAAGCAACGCGGCTTCTGTTGTCACCACTCTGAGCAAC 900
901 GATCGGTTAGTGGGATTTGGGAACAAGAGCTGATGATGCGCCAGCGTATTCAGCGT 960
901 GATCGGTTAGTGGGATTTGGGAACAAGAGCTGATGATGCGCCAGCGTATTCAGCGT 960
961 ATGCGTCAGTTGTTTCTGCTCAATACGCTGAGGAAAGCGCAACCGCGACTTCAGCTTT 1020
961 ATGCGTCAGTTGTTTCTGCTCAATACGCTGAGGAAAGCGCAACCGCGACTTCAGCTTT 1020
1021 ATCATCAACAGAACGCGATGTTCTCTTTCAGTGGCCTGACAAAGAACCAAGTCTCGT 1080
1021 ATCATCAACAGAACGCGATGTTCTCTTTCAGTGGCCTGACAAAGAACCAAGTCTCGT 1080
1081 CTGCGCGAAGATTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTCGCGGCGATG 1140
1081 CTGCGCGAAGATTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTCGCGGCGATG 1140
1141 ACACAGATACATGGCTCGCTGCGAAGCGATTTGGCAGTGTGTA 1191
1141 ACACAGATACATGGCTCGCTGCGAAGCGATTTGGCAGTGTGTA 1191

```

RESULT 2

```

US-10-673-786A-1
; Sequence 1, Application US/10673786A
; Publication No. US20040132165A1
; GENERAL INFORMATION:
; APPLICANT: AKHVERDIAN, VALERY ZAVENOVICH
; APPLICANT: SAVRASOVA, EKATERINA ALEKSEEVNA
; APPLICANT: KAPLAN, ALLA MARKOVNA
; APPLICANT: LOBANOV, ANDREY OLEGOVICH
; APPLICANT: KOZLOV, YURI IVANOVICH

```

```

; TITLE OF INVENTION: METHOD FOR PRODUCING L-THREONINE USING BACTERIA
; FILE OF INVENTION: BELONGING TO THE GENUS ESCHERICHIA
; CURRENT APPLICATION NUMBER: US/10/673,786A
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/JP03/02067
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: 2002104983
; PRIOR FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1191)
US-10-673-786A-1

Query Match      100.0%; Score 1191; DB 7; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGGAGAACATTACCGCGCTCTGCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 1 ATGTTTGGAGAACATTACCGCGCTCTGCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Qy 61 CGTGCCGATGAACGCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 61 CGTGCCGATGAACGCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Qy 121 GGCAAAACCCCGGTACTGACAGCGTGAAGAGCTGAACAGTATCTGCTCGCAAAATGAA 180
Db 121 GGCAAAACCCCGGTACTGACAGCGTGAAGAGCTGAACAGTATCTGCTCGCAAAATGAA 180
Qy 181 ACCACAAAATTTACCTCGGCATTTGACGCGATTTGCGTGAATTTGGTGTCTATAAAGATGAGACG 240
Db 181 ACCACAAAATTTACCTCGGCATTTGACGCGATTTGCGTGAATTTGGTGTCTATAAAGATGAGACG 240
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGCGCACAGACT 300
Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGCGCACAGACT 300
Qy 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACAGCGTT 360
Db 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACAGCGTT 360
Qy 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Qy 421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATCGGAAATACACTCTTGAATTCGAT 480
Db 421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATCGGAAATACACTCTTGAATTCGAT 480
Qy 481 GCATCTGATTTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 540
Db 481 GCATCTGATTTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 540
Qy 541 TGCCATTAACCAACCCGCTATCGCCCTACGCTGGAACAATGGCAACACTGGCACAACTC 600
Db 541 TGCCATTAACCAACCCGCTATCGCCCTACGCTGGAACAATGGCAACACTGGCACAACTC 600
Qy 601 TCCGTTGAGAAAGGCTGTGTTACCGCTGTTGACTTTCGCTTACCAGGGTTTTGGCCGCGGT 660
Db 601 TCCGTTGAGAAAGGCTGTGTTACCGCTGTTGACTTTCGCTTACCAGGGTTTTGGCCGCGGT 660
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGCTATGCAATAAAGAGCTGATTTGTT 720
Db 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTCGCGCTATGCAATAAAGAGCTGATTTGTT 720
Qy 721 GCCAGTTCCTACTCTAAACCTTTGGGCTGTGACAGAGCGTGTGGCGCTTGTACTCTG 780

```


Db 721 GCCAGTTCTTCTATAAACTTTGGCTGTACACGAGCGTGTGGCGTTGTACTCTG 780
Qy 781 GTTGCTCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 840
Db 781 GTTGCTCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 840
Qy 841 GCTAACTACTCTAACCCACGACACACGCGGCTTCTGTTGTCACCATCTCAGCAAC 900
Db 841 GCTAACTACTCTAACCCACGACACACGCGGCTTCTGTTGTCACCATCTCAGCAAC 900
Qy 901 GATGCGTTAGTGGATTTGGGAACAAGAGTGAATGAGCCAGCGTATTCAGCGT 960
Db 901 GATGCGTTAGTGGATTTGGGAACAAGAGTGAATGAGCCAGCGTATTCAGCGT 960
Qy 961 ATGCGTCAGTTGTTGTCGTCATAGCTGACGGAAGGCGCAACCGCGACTTCAGCTT 1020
Db 961 ATGCGTCAGTTGTTGTCGTCATAGCTGACGGAAGGCGCAACCGCGACTTCAGCTT 1020
Qy 1021 ATCATCAAAACAGAACGCGATGTTCTCTTTCAGTGGCCTGACAAAGAACCAAGTCTGCGT 1080
Db 1021 ATCATCAAAACAGAACGCGATGTTCTCTTTCAGTGGCCTGACAAAGAACCAAGTCTGCGT 1080
Qy 1081 CTGCGGAAAGATTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCGGATG 1140
Db 1081 CTGCGGAAAGATTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCGGATG 1140
Qy 1141 ACACCAATACATGCTCCGCTGTGCGAAGCGATTTGTGCGAGTGTGTA 1191
Db 1141 ACACCAATACATGCTCCGCTGTGCGAAGCGATTTGTGCGAGTGTGTA 1191

RESULT 3

US-10-893-671-87/c
; Sequence 87, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PKZ-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 14759
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-893-671-87

Query Match 100.0%; Score 1191; DB 9; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTTGAGAACATTAACCGCGCTCTGCGGACCCGATTCGGGCTCGCGGATCTGTTT 60
Db 2755 ATGTTTGAGAACATTAACCGCGCTCTGCGGACCCGATTCGGGCTCGCGGATCTGTTT 2696
Qy 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTCTATAAGATGAGACG 120
Db 2695 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTCTATAAGATGAGACG 2636
Qy 121 GGCAAAACCCGGTACTGACAGCGTGAAGAGCTGAAGAGCTGATCTGCTCGAAATGAA 180
Db 2635 GGCAAAACCCGGTACTGACAGCGTGAAGAGCTGAAGAGCTGATCTGCTCGAAATGAA 2576
Qy 181 ACCACCAAAATTAACCTCGGATTTGAGCGGATCCCTGAAATTTGGTCTGCTCACTCAGGAA 240
Db 2575 ACCACCAAAATTAACCTCGGATTTGAGCGGATCCCTGAAATTTGGTCTGCTCACTCAGGAA 2516

Qy 241 CTGCTCTTTTGGTAAAGGTAGCGCCTGATCAATGACAAACGTCGTCTGCGACGACAGACT 300
Db 2515 CTGCTCTTTTGGTAAAGGTAGCGCCTGATCAATGACAAACGTCGTCTGCGACGACAGACT 2456
Qy 301 CCGGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTGCGCAAAATAACAGCGTT 360
Db 2455 CCGGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTGCGCAAAATAACAGCGTT 2396
Qy 361 AACGCGTGTGGGTGAGCAACCCAAAGCTGCCCAACCATAGAGCGTCTTTAACTCTGCA 420
Db 2395 AACGCGTGTGGGTGAGCAACCCAAAGCTGCCCAACCATAGAGCGTCTTTAACTCTGCA 2336
Qy 421 GGTCTGGAAGTTTCTGTAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
Db 2335 GGTCTGGAAGTTTCTGTAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 2276
Qy 481 GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 540
Db 2275 GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 2216
Qy 541 TGCATTAACCCCAACCGGTATCGACCTAGCTGGAACAAATGGAACACACTGGCAAACTC 600
Db 2215 TGCATTAACCCCAACCGGTATCGACCTAGCTGGAACAAATGGAACACACTGGCAAACTC 2156
Qy 601 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGCTTTTGGCCGCTGGT 660
Db 2155 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGCTTTTGGCCGCTGGT 2096
Qy 661 CTGGAAGAAGATCTGAAAGACTGCGGCTTTTCCGCGCTATGATGCAATGAAGAGTGAATGTT 720
Db 2095 CTGGAAGAAGATCTGAAAGACTGCGGCTTTTCCGCGCTATGATGCAATGAAGAGTGAATGTT 2036
Qy 721 GCCAGTTCTTACTCTAAAGACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
Db 2035 GCCAGTTCTTACTCTAAAGACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 1976
Qy 781 GTTGCTGCGGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 840
Db 1975 GTTGCTGCGGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAAGCGGATTCGC 1916
Qy 841 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTTGGCCACCATCTCTGAGCAAC 900
Db 1915 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTTGGCCACCATCTCTGAGCAAC 1856
Qy 901 GATGCGTTAGTGGATTTGGGAACAAGAGCTGATATGCGCAGCGTATTCAGCGT 960
Db 1855 GATGCGTTAGTGGATTTGGGAACAAGAGCTGATATGCGCAGCGTATTCAGCGT 1796
Qy 961 ATGCGTCAGTTGTTCTGTCATAGCTGAGGAAAAAGGCGCAAAACCGCGACTTTCAGCTTT 1020
Db 1795 ATGCGTCAGTTGTTCTGTCATAGCTGAGGAAAAAGGCGCAAAACCGCGACTTTCAGCTTT 1736
Qy 1021 ATCATCAAAACAGAACCGCATGTTCTCTTTCAGTGGCTGACAAAGAACCAAGTCTCGCT 1080
Db 1735 ATCATCAAAACAGAACCGCATGTTCTCTTTCAGTGGCTGACAAAGAACCAAGTCTCGCT 1676
Qy 1081 CTGCGGAAAGATTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCGGATG 1140
Db 1675 CTGCGGAAAGATTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCGGATG 1616
Qy 1141 ACACCAATACATGCTCCGCTGTGCGAAGCGATTTGTGCGAGTGTGTA 1191
Db 1615 ACACCAATACATGCTCCGCTGTGCGAAGCGATTTGTGCGAGTGTGTA 1565

RESULT 4

US-10-369-493-44812
; Sequence 44812, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 44812
 ; LENGTH: 1188
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus nematophilus
 US-10-369-493-44812

Query Match 55.1%; Score 655.8; DB 6; Length 1188;
 Best Local Similarity 72.0%; Pred. No. 3.8e-216;
 Matches 855; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

Qy	1	ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATCTGGGCGCTGGCGGATCTGTTT	60
Db	1	ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATCTGGGCGCTGGCGGATCTGTTT	60
Qy	61	CGTGCCGATGACGCTCCCGGCAAAATTAACCTCGGATTTGGTCTCTATAAGATGAGACG	120
Db	61	AAAGCTGATCTCTGTAATAAATAAATCAATCTGGGTATCGGTGTTTACAAAGACGAACT	120
Qy	121	GGCAAAACCCGGTACTGACGAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA	180
Db	121	GGTAAACCCCTGCTGACGACCGTTAAAGAGCAGAAAATCTGCTGGAAACGAA	180
Qy	181	ACCACCAAAATTAACCTGCGCATTCGCGGATCCCTGAATTTGGTCTGCTCCACTCAGGAA	240
Db	181	ACCACCAAAATTAACCTGCGCATTCGCGGATCCCTGAATTTGGTCTGCTCCACTCAGGAA	240
Qy	241	CTGCTGTTGGTAAAGTAGGCGCTGATCAATGACAAAGCTGCTGCGACGACAGACT	300
Db	241	CTACTCTTCGCAATACGACGACTATCAATGACAAAGCTGCTGCGACGACAGACT	300
Qy	301	CCGGGGGCACTGCGGCACTACGCTGCGTGGCTGCGATTTCTGCGCAAAATACCAAGGTT	360
Db	301	CCAGCGGCAACCGGTGCACTGCGGCTGCGGAGATTTATGCTAAACAGACCAATGCA	360
Qy	361	AAGCGTGTGGTGAGCAACCCCAAGCTGCGGCAACCAATGAGAGCGCTTTAACTCTGCA	420
Db	361	AAACGTGTGATCAGTAACCCCAACATGCGCAACCAATGAGAGCGCTTTCTCTAGCGCT	420
Qy	421	GGTCTGGAAGTTCTGTAATAGCTTTATATGATGCGGAAATCACTCTTGACTTCGAT	480
Db	421	GGGTAGAAATCCCGAGTAGTACCAATATTAACAATGAGAAAGCATGCTCTGGATTTTAT	480
Qy	481	GCACGTGATTAACAGCTGAATGAGCTCAGCTGCGGACGCTAGTGTGTTCCATGCTGC	540
Db	481	GGCATGTGTGCTAGCTGCTGAAAGCTCAAGCAGGTATGTTGTTCTGCTTACGGTTGC	540
Qy	541	TGCCATAACCCCAACCGGTATCGATCCAAACCGCTGAACAGTGGCAGAACTGGCAGATTA	600
Db	541	TGCCATAACCCCAACCGGTATCGATCCAAACCGCTGAACAGTGGCAGAACTGGCAGATTA	600
Qy	601	TCCGTTGAGAAAGCTGGTACCGCTGTTTGAATTCGCTTACCAAGGTTTGGCCGCGGT	660
Db	601	TCTCGGCAAAATGGCTGGCTGCTGTTTGAATTCGCTTATCAGGTTTGGCCAGAGT	660
Qy	661	CTGGAAGAGATGCTGAAGGACTGCGGCTTTTCGCGCTATGATCAAGAGCTGATGTT	720
Db	661	TTAGATGAAGATGAGAGGCTCGGTATTTTACTAAACCAATATGAATGAATGATCGTT	720
Qy	721	GCCAGTTCTTACTTAAACCTTTGGGCTGTGTAACGAGCGGTGTTGGCGTTTGTACTCTG	780
Db	721	GCCAGCTCTTACTCAAAAATTTTGGGCTGTGTAACGAGCGGTGTTGGCGTTTGTACTCTG	780

Qy	781	GTTGTGCGCAGTGAACCCGTTGATCGCATTCAGCCAAATGAAGCGCGATTCG	840
Db	781	GTTGCAACTGACAGCGATACCCGAGAAAAGCATTTCAGTCAGCGAAATCTATTGTTCTG	840
Qy	841	GCTAACTACTTAACCCACAGACACGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	900
Db	841	ACTAACTACTTAACCCACAGCGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	900
Qy	901	GATCGTTACGTCGATTTGGGNAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT	960
Db	901	GACGAATTTAAAGCAGAAATGGAATTCAGAACTGCGGACTATGCGTGAACGCAATTCGTCG	960
Qy	961	ATGCGTCAGTTGTTGTCATATACGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1020
Db	961	ATGCGTCAGTTGTTGTCATATACGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1020
Qy	1021	ATCATCAAAACAGAACCGGATGTTCTCTTCAGTGGCTGACAAAGAAACAAAGTGTGCGGT	1080
Db	1021	ATTATTAGCAAAATCGTATGTTTTCATTTCAGTGGTCTGACTAAAGAAACAAAGTGTGATCG	1080
Qy	1081	CTGCGGAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGTAAATGTTGCGCGGATG	1140
Db	1081	CTGCGTGAAGAATTCGTTATTTATGCGGTGAGTCTGCGGCGCAATTAACGTAGCGGCTTA	1140
Qy	1141	ACACAGATAACATGCTCCGCTGTCGAAAGGAGTGTGCGCAGTGCT	1187
Db	1141	ACACTGAGAACATGTTCTCTGTTGAGCGATTTGTTGCACTACT	1187

RESULT 5

US-10-369-493-23918
 ; Sequence 23918, Application US/10369493
 ; Publication No. US2003023675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23918
 ; LENGTH: 966
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus nematophilus
 US-10-369-493-23918

Query Match 42.8%; Score 510; DB 6; Length 966;
 Best Local Similarity 70.5%; Pred. No. 1.6e-165;
 Matches 681; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

Qy	223	GCTGCTGCACTCAGGAATGCTGTTTGGTAAAGGTAGCGCTTCAATCAATGCAACCGT	282
Db	1	GGCCGGGTGATTTCAGGAATGCTGTTTGGAAAGTACCAGTGCCATCTTACTGAAACCGT	60
Qy	283	GCTGCGACGGCAGACTCGGGGGGCACTGGGGGCACTAGCGGTGCTGCCGATTTCTTG	342
Db	61	GCCCGTACCGTACAAAGCCCGGAGCAGCGGCACTGCGTACTGCGCTGATTTTATT	120
Qy	343	GCAAAAAATACCAAGCGTTAAGCGTGTGGTGAGCAACCCAAAGCTGGCGAACCAATAAG	402
Db	121	GCCACCAAACTATGCCAAACGTGTTGGATCAGTAACCCAACTGGCCAAACCAATAA	180
Qy	403	AGCGTCTTTAACTCTGAGGTCTGGAAGTTCGTGAATACGCTTTATGATCGGGAAT	462
Db	181	GGTGTGTTTTTCCAGCGCAGGTTTAGAGATCCCGGAATATAACTATTACGATGAGAAAA	240

```
Qy 463 CACACTCTTGACTTCGATGCACTGAATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGTA 522
Db 241 CACGCTCTGAATTTGCAAGGCGATCTGGCAAGCTGTCTGAAGCACAAAGCGGCGTATGTT 300
Qy 523 GTGCTGTTCATGGCTGCTGCCATAACCCCAACCGGTATCGACCTACGCTGGGAACATGG 582
Db 301 GTTCTGTTCACGGTTCCTGCCATAACCCCAACCGGTATCGATCTCTCCAGAACAGTGG 360
Qy 583 CAAACACTGGCACAACCTCTCCGTTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTAC 642
Db 361 CAAAACCTGGCAGATTTATCTGGGCAATAGGCTGGCTGCTGTATTCGATTTGCTTAT 420
Qy 643 CAGGTTTTGCGCGTGTCTGGAAGAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATG 702
Db 421 CAGGGCTTTGCTAAAGTCTGGATGAAGATGCGGAAGGCTCGGCGATTTTTCGCAAAAAC 480
Qy 703 CATAAAGAGCTGATTTGTCAGTTCTTACTCTTAAAACTTTGGCCTGTACAAAGCGGT 762
Db 481 CATAATGAATTGATCGTTGCGAGCTCTTACTTCAAAAACTTTGGCCTGTACAAATGAGCGT 540
Qy 763 GTTGGCGCTTGTACTCTGGTTGTCGCGACAGTGAACCGTTGATCGCGATTCAGCCAA 822
Db 541 GTTGGTCCCTGTACTATTTGTCAGCAGACAGGATACCGAGAAAGGTTTAGTCAG 600
Qy 823 ATGAAGCGCGGATTCGCGCTAACTACTTAACCCACGACACACGCGCTTCTCTGTTGT 882
Db 601 GCTAAATTTATGTACGTACCAACTATTTCAACCCACGCGCTCATGGTCTTCTGTTGT 660
Qy 883 GCACCATCTGAGCAACGATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATG 942
Db 661 ACACAAATTTCTGCCAAAGAACTTGAAGCAGAGTGGATTCAGGAACCTGGCAACAATG 720
Qy 943 CGCAGCGTATTCAGCGTATGCGTTCAGTTGTTGTCATACGCTGCAGGAAAGGCGCA 1002
Db 721 CGTAGCGCATTCAGCGTATGCGCAATTTGTTGTCACACCTTACAGAAAGAGTGCA 780
Qy 1003 AACCGGACTTCAGCTTTATCATCAACAGACGGCATGTTCTCTTCAGTGGCCTGACA 1062
Db 781 AAACAGGATTTACGTTTCATCTGTGCTCAAAATGGTATGTTCTCAATTCAGCGGTCTGACT 840
Qy 1063 AAGAACAGTGTGCGTCTGCGGAAGAGTTTGGCGTATATGGGTTGCTTCTGTCGC 1122
Db 841 AAGAACAGATTTGAACGCTTCGCGTGAAGATACGCTGTTTATGCAAGTTCAGCTCCGCGCT 900
Qy 1123 GTAATGTGCGCGATGACACAGATAACATGCTCCGCTGTGCGAAGCGATTGGCA 1182
Db 901 ATTAACGTCGCTGTTTAACTGGAACATGTTCTCTGTGTGAAGCGATTGTTGCT 960
Qy 1183 GTGCTG 1188
Db 961 GTACTG 966
```

```
RESULT 6
US-10-450-763-23995
; Sequence 23995, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-31
; PRIOR FILING DATE: 2000-03-31
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23995
; LENGTH: 2133
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)-(774)
; OTHER INFORMATION: 90% homologous to Escherichia coli ATP-dependent helicase
; OTHER INFORMATION: HrpA homolog, accession number D90780, Smith-Waterman Score=1183.
US-10-450-763-23995

Query Match 41.7%; Score 496.2; DB 9; Length 2133;
Best Local Similarity 99.4%; Pred. No. 1.5e-160;
Matches 498; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 687 CGCTTTTCGCGCTATGCATAAAGAGCTGATTTGTTGCCAGTTCTTACTCTTAAACCTTTGG 746
Db 1425 CCCTTTTCGCGCTATGCATAAAGAGCTGATTTGTTGCCAGTTCTTACTCTTAAACCTTTGG 1484
Qy 747 CTTGTACAACGAGCGGTGTTGGCGCTTGTACTCTGTTGCTGCCGACAGTGAACCCGTTGA 806
Db 1485 CTTGTACAACGAGCGGTGTTGGCGCTTGTACTCTGTTGCTGCCGACAGTGAACCCGTTGA 1544
Qy 807 TCSCGCATTTCAGCCAAATGAAGCGCGATTTCGCGCTAACTACTCTAAACCCACGACACA 866
Db 1545 TCSCGCATTTCAGCCAAATGAAGCGCGATTTCGCGCTAACTACTCTAAACCCACGACACA 1604
Qy 867 CGCGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACGATGCGTTACGTCGCAATTTGGGAACA 926
Db 1605 CGCGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACGATGCGTTACGTCGCAATTTGGGAACA 1664
Qy 927 AGAGCTGACTGATATGCGCGAGCGTATTCAGCGTATGCGTTCAGTTGTTGCTCAATAGCT 986
Db 1665 AGAGCTGACTGATATGCGCGAGCGTATTCAGCGTATGCGTTCAGTTGTTGCTCAATAGCT 1724
Qy 987 GCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACGCGCATGTTCTC 1046
Db 1725 GCATGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACGCGCATGTTCTC 1784
Qy 1047 CTTTCAGTGGCTGCACAAAAGAACAAAGTGTGCTGCGCAAGAGTTTGGCGTATATGC 1106
Db 1785 CTTTCAGTGGCTGCACAAAAGAACAAAGTGTGCTGCGCAAGAGTTTGGCGTATATGC 1844
Qy 1107 GTTTGTCTTCTGTCGCGTAAATGTTGCGCGGATGACACAGATAAATGCGTCCGCTGTG 1166
Db 1845 GTTTGTCTTCTGTCGCGTAAATGTTGCGCGGATGACACAGATAAATGCGTCCGCTGTG 1904
Qy 1167 CGAAGCGATTGTCGAGTGTCT 1187
Db 1905 CGAAGCGATTGTCGAGTGTCT 1925
```

```
RESULT 7
US-10-450-763-25886
; Sequence 25886, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 25886
; LENGTH: 3222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

; NAME/KEY: SIMILAR
; LOCATION: (1)..(774)
; OTHER INFORMATION: 90% homologous to Escherichia coli ATP-dependent helicase
; OTHER INFORMATION: hrpa homolog, accession number D90780, Smith-Waterman Score=1183.
US-10-450-763-25886

Query Match 41.7%; Score 496.2; DB 9; Length 3222;
Best Local Similarity 99.4%; Pred. No. 1.9e-160; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 3;

Qy 687 CGCTTTCCGGCTATGCATGAAGAGCTGATTTGTTGCCAGTTCCTACTCTATAAACTTTGG 746
Db 1425 CCCTTTCGGCGCTATGCATGAAGAGCTGATTTGTTGCCAGTTCCTACTCTATAAACTTTGG 1484
Qy 747 CTTGTACAAACGAGCGTGTGGCGTGTGTAATCTCTGGTTGTCGCCGACAGTGAACCGTTGA 806
Db 1485 CCTGTACAAACGAGCGTGTGGCGTGTGTAATCTCTGGTTGTCGCCGACAGTGAACCGTTGA 1544
Qy 807 TCGCGCATTCAGCCCAATGAAGCGCGATTTCGGCGTAATCTACTCTAACCCACGACACA 866
Db 1545 TCGCGCATTCAGCCCAATGAAGCGCGATTTCGGCGTAATCTACTCTAACCCACGACACA 1604
Qy 867 CGGCGCTTCGTGTTGCCACCATCTCTGAGCAACGATGCGTTACGTGCGATTTTGGGAACA 926
Db 1605 CGGCGCTTCGTGTTGCCACCATCTCTGAGCAACGATGCGTTACGTGCGATTTTGGGAACA 1664
Qy 927 AGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCAGTTGTCGTAATAGCGT 986
Db 1665 AGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCAGTTGTCGTAATAGCGT 1724
Qy 987 GCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACCGCATGTTCTC 1046
Db 1725 GCATGAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACCGCATGTTCTC 1784
Qy 1047 CTTTCAGTGGCTGACAAAGAAAGAGTGTGCGTCTGGCGAAGATTTGCGGTATATGC 1106
Db 1785 CTTTCAGTGGCTGACAAAGAAAGAGTGTGCGTCTGGCGAAGATTTGCGGTATATGC 1844
Qy 1107 GGTTCGTTCTGTCGCTGAAATGTGGCGGAGTACACCGATGAATGCTCGGCTGTG 1166
Db 1845 GGTTCGTTCTGTCGCTGAAATGTGGCGGAGTACACCGATGAATGCTCGGCTGTG 1904
Qy 1167 CGAAGCGATTTGCGAGTGCT 1187
Db 1905 CGAAGCGATTTGCGAGTGCT 1925

RESULT 8
US-10-795-159-536
; Sequence 536, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 536
; LENGTH: 7977
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4108)..(4108)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7960)..(7960)
; OTHER INFORMATION: n = a, c, g, or t

US-10-795-159-536
Query Match 41.4%; Score 493.4; DB 9; Length 7977;
Best Local Similarity 63.4%; Pred. No. 3.1e-159; Indels 0; Gaps 0;
Matches 755; Conservative 0; Mismatches 436;

Qy 1 ATGTTTGGAGAACATTACCGCGCTCTGCGGACCCCGATCTTGGCCCTGCGCGATCTGTTT 60
Db 1594 ATGTTTGAACATATTAAAGCGGACCGATCCATCTTAGGCTTAGCGGAAGCATTC 1653
Qy 61 CGTGGCGATGAACGTCCTCCGCGCAAAATTAACCTCGGAGTTGGTGTCTATATAAGATGAGACG 120
Db 1654 ARAATCCGAAACTCGCGAAATATAATTAACCTGGGTATTGGCGTTTATAAGATGCGCAA 1713
Qy 121 GGCAGAAACCCCGTACTGACAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 180
Db 1714 GGCACAAACCCCGATTTATGCGTGAAGAGGAGAGAGAAACCGCTATTATTGATAGGAA 1773
Qy 181 ACCACCAAAATTTACCTCGGCATTGACGCGATCCCTGAAATTTTGGTCGCTCAGGAA 240
Db 1774 AMAACGAAATTTATCTGACTATCGATGGTATTGCTGATTAACGAAACAAACGACG 1833
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCTCGACGGCACAGACT 300
Db 1834 TTACTTTTTCGGTAAAGATTCTGAAGTCATCAAAATCTAATCGTCAAGAACAGTACAAAGT 1893
Qy 301 CGGGGGGCGACTGGCGCAGTACGCGTGGCTGCGGATTTCTTGGCAAAATATACGAGGTT 360
Db 1894 TTAGCGGGAACAGCGGCATTTACGCAATTGGCGGAGAAATTTATAACGCTCAAACTAAGCA 1953
Qy 361 AAGCGTGTGCGGTGAGCAACCAACGCTGCGGCAACCAATGAAGAGCGTCTTTAACTCTGCA 420
Db 1954 CAAATGTTGGATAGCACCCCACTTGGCCAAACCAATGCCATTTTCAATGCGCTC 2013
Qy 421 GGTCTCGAAGTTGCGTAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 480
Db 2014 GGTATGCAATTCGTGAATATCGTTATTATGATGCTGACGCAAGCCCTTGTATGGAT 2073
Qy 481 GCATGATTAAACGCTGAAATGAAGCTCAGGCTGCGGACGATGCTGTGTTCCATGCGTGC 540
Db 2074 CATTTACTTGAAGATTAAAGCCAAAGCAAGGCGATGTGTTGCTTTTACACGGTGT 2133
Qy 541 TGCCATAACCCAAACCGGTATCGACCTAGCTGGAACAATGGCAACACACTGGCACAACTC 600
Db 2134 TGCCATAATCCGACTGGTATGACCCCTACTCCAGAAACAATGGCAAGAAATAGCCGCACTT 2193
Qy 601 TCGTTTGAGAAAGCGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGGCCGTTGT 660
Db 2194 TCGGCGAAACCGCTGGTTGCCACTTTTGACTTTTGCTTATCAAGTTTAGCCAATGGA 2253
Qy 661 CTGGAAGAGATGCTGAAGGACCTGCGCGCTTTTGGCGGCTATGATTAAGAGCTGATTGTT 720
Db 2254 TTAGATGAAGATGCTTATGGCTTGGCTGCTTTTGGCGCAACCCACAAAGAAATTTAGTA 2313
Qy 721 GCCAGTTCCTACTCTAAACACTTTGGCTGTACACAGGCGTGTGCGGCTTCTACTCTG 780
Db 2314 GCGAGTTCAATCTCGAAACACTTTGGTTTATTAATGAACGCTGTTGGTCTTTACTCTT 2373
Qy 781 GTTGTGCGCGACAGTGAACCGTTGATCGCGCAATTCAGCCAAATGAAGCGGCGATTTCG 840
Db 2374 GTGGCAGAAATGACAGAAATTCATCAACCGCATTAACCCAGTGAATCAATATTTCGT 2433
Qy 841 GCTAACTACTCTAACCCACGACACAGCGGCTTCTGTTGTTGCGCCACCATCTCTGAGCAAC 900
Db 2434 ACATCTCTACTTAACCCCTGCATCTCACGGCGGCGCAACTGTAGCAACAGATTATAATGAC 2493
Qy 901 GATGCGTTAGTGGGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Db 2494 ACTCAACTTCGCCAAGAGTGGGAAATGAATTAACCTGAATTCGCGAAACGATTCACAAAA 2553
Qy 961 ATGCGTCAGTTGTTGCTCAATACGCTGAGGAAAAAGGCGCAACCGCGACTTCAGCTTT 1020
Db 2554 ATGCGTCATTTATTCGTTTCAAGTTTAAAGAAATACGGTTCGGAACAAGATTTCAGCTTT 2613


```

; FILE REFERENCE: GJE-6436
; CURRENT APPLICATION NUMBER: US/10/275,026A
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: PCT/GB01/02003
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: GB 0011108.8
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 187
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1191)
US-10-275-026A-187

Query Match          39.6%; Score 471.8; DB 7; Length 1191;
Best Local Similarity 62.3%; Pred. No. 3.3e-152;
Matches 740; Conservative 0; Mismatches 447; Indels 0; Gaps 0;

QY 2 TGTTTGAGACATTACCGCGCTCCTGCCGACCGGATTTGGGCGCTGGCGGATCTGTTTC 61
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 5 TCTTCAAGCACATCGAAGCGCGCGCGCGCATCGGATCTCGGTTTGGGCGAGCATTC 64
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 62 GTCCGAGTAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACGG 121
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 AAGCGAAACCCCGCGGAAAGTCAACCTCGGCATCGCGGTGTACAAAGAGCGCATCCG 124
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 GCAAAACCCGGTACTACGACGAGTGAAGAGCTGACAGATATCTCTCGAAATGAAA 181
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 GCGCGACACCTATTGTCAAGCGCTCAAGAGCGCGGAAACCGCTGTTGGAAAGCGAAA 184
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 CCACCAAAATTTACCTCGCATTCAGCGCATCCTGAAATTTGGTGTGCTGCACTCAGGAAC 241
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 185 CAATAAAATACCTCACCATCGACGCGTGGCGACTACACGACCAACCCAAATTC 244
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 242 TGCTGTTTGGTAAGTAGCGGCTGATCAATGACAAACGCTGTGCGACGGGACAGACTC 301
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 245 TGCTGTTTGGCAAGACCAAGCAATCATCGCAGCGCTGCGGCGCAAAACAGCGCAAGCC 304
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 302 CGGGGGGCACTGGGCACTAGCGGTGCTGCGGATTTCTGGCAAAATACAGCGTTA 361
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 305 TTGCGGTACAGGCGCATTCGATTTGCGGCGGAGTTTGCCAAACGTCAGTGAACGCGC 364
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 362 AGCGTGTGGGTGAGCAACCCAGCTGGCGGCAACCAAGAGCGTCTTTAACTCTGCGAG 421
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 365 AAACCATCTGGATTTCCATCCGACTTGGGCCCAACCAACGCGCATGCCAAAGCGGTG 424
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 422 GTCTGGAAGTTTGGTGAATAGCGTTATTATGATCGGGAATATCACTCTTTGACTTCGATG 481
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 425 GTATCCAAAGACCAACCTTATCGCTTACTATGATGCGCGCAAAACAGGTTTGGATTGGGACG 484
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 482 CACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTTCATGGCTGCT 541
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 485 GTATGATTGAAGACTTGAAGCAAGGCAAAAGGCGCATCGTCTCTGTCGACGCGTGTCT 544
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 542 GCCATACCCCAACCGGTATGACCCCTACGCTGGAACAATGCAAACTGCGCAACTCT 601
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 545 GCCCAACCCCTACCGGTATGACCCCTACGCGCGCAACAAATGGAAACTCTGGCAAACTTT 604
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 602 CCGTTGAGAAGGCTGTATTACCGCTGTTTGAATTCGCTTACCAGGGGTTTTGCCCGTGGTC 661
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 605 CTGCGGAAAGGCTGTATTACCGCTGTTTGAATTCGCTTACCAGGGCTTCGGCAATGTT 664
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 662 TGAAGAAGATGCTGAAGAGACTGCGGCTTTGCGGCTATGCATATAAGAGCTGATTTG 721
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 665 TGAAGAAGATGCTGACGCGCTGCGGCTGTTCTTTGAAACACAATACAGAAATTTGCTGATTG 724
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 722 CCAATCTCTCTAAAACTTTGGGCTGTACACAGAGCGGTGTGGCGCTTGACTCTGG 781
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 725 CCAGCTCTTATCCAAAACCTTCGTTATGTACACAGAGCGGTGCGGCGGTTTCACTTTGG 784
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 11
US-10-915-740A-40
; Sequence 40, Application US/10915740A
; Publication No. US20050191316A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tettelin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scarliato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; CURRENT FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 17381
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-40

Query Match          39.2%; Score 466.8; DB 9; Length 17381;
Best Local Similarity 62.0%; Pred. No. 8.3e-150;
Matches 738; Conservative 0; Mismatches 452; Indels 0; Gaps 0;

```


Db	562674	TCGGCGGTACGGGCGCATTCGGTATTGCGCGCGAGTTTGCCAAACGCCAGTTGAACGCGC	562615
Qy	362	AGCGTGTGGGTGAGCAACCCAGCTGGCGCGAACATTAAGAGCGTCTTTAACTCTGCAG	421
Db	562614	AAACCATCTGGATTTCCAAATCCGACTTTGGCCCAACCAACGCGCATCGCCAAAGCGGTGC	562555
Qy	422	GTCTGGAAGTTCGTGAATACGCTTATTATGATCGGGAATACACATCTTGACTTCGATG	481
Db	562554	GTATCCAGACAACCTTATCTGTTACTATGATGCCGCAACACGCTTTGGATTTGGGACG	562495
Qy	482	CACGTATTAACAGCTCGAATGAAGCTCAGGCTGGCGACGTAAGTCTGCTTCCATGGCTGCT	541
Db	562494	GCATGTTGAGGACTTGGCCAGCGCGCAAAAGGCGACATCGTCTGCTGCACGGTGTCT	562435
Qy	542	GCCATTAACCAACCGGTATCGACCTTACGCTGCGTGAACAAATGGCAAAACACTGGCAACTCT	601
Db	562434	GCCACAATCTACCGGCTACGACCTTACGCGCGCAACAAATGGGAACCTTTGGCAAACTTT	562375
Qy	602	CCGTTGGAAGAGCTGGTTACCGCTGTTTGACTTTCCTTACCAAGGCTTTGCCCGTGGTC	661
Db	562374	CTGCCGAAAAGGCTGGTTGCCGCTGTTTGACTTTCCTTACCAAGGCTTCGGCAATGGTT	562315
Qy	662	TGGAAGAAGATGCTGAAGGACTGCGGCTTTTCGGGCTATGCATAAAGAGCTGATTGTTG	721
Db	562314	TGGAGNAGATGCTACGGCTTGGCGGTGTTCTTGAAACAAATACAGATTGCTGATG	562255
Qy	722	CCAGTTCCTACTCTAAACCTTTTGGCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG	781
Db	562254	CCAGCTCTTATTCAAAACCTTCGGTGTGTACAAACGAGCGGTGGCGGCTTCACCTTTGG	562195
Qy	782	TTGCTGCGCAGTGAACCGTGTGATCGCGCATTCAGCCAAATGAAGCGGCGATTCGCG	841
Db	562194	TGGCCGAAGATGAAGAAACAGCAGCGCGCCGCCACAGCCAAAGTCAAAACCATCATCCGT	562135
Qy	842	CTAACTACTCTAACCCACAGCAGCAGCGGCGCTTCTGTTGTGCCACCATCTCAGCAACG	901
Db	562134	CTTTGTATTCCAACCCGGCTTCAACGCTGCGAACCATTTGCGTGTGTTGAAAATG	562075
Qy	902	ATGCGTTPACGTGGATTTTGGGAACAAGAGCTGATGATATCGCCAGCGTATTTCAGCGTA	961
Db	562074	ATGATTGAAGCACAATGATTGCCGAATCGATGAATGCGGCGCGCATCAAAAGCCA	562015
Qy	962	TGGCTCATGTTGTCGTCATACGCTGCAGGAAAGGCGCAAAACCGGACTTCAGCTTTTA	1021
Db	562014	TGGCGCAAAATTTGTCGGGTGCTCAAAGCCAAAGGTGCAAGCCAAACCTTTGATTTC	561955
Qy	1022	TCATCAACAGAGCGGATGTTCTCTTCAGTGGCCCTGACAAAGAACAGTCTCGCTC	1081
Db	561954	TTATCAACAAACGGTATGTTCTTTTCAGCGGCTTGACTCCCGCAACAGTCCGCGCC	561895
Qy	1082	TGGCGGAAGATTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGGCCGGATGA	1141
Db	561894	TGAAAACGAGTTTGCAATTTATGCGCTCGCTCCGCGCGCATCAACGTCGCGGCATTA	561835
Qy	1142	CACAGATAACATGGCTCCGCTGTGCGAAGCGATTTGGCAGTGTCTGTAA	1191
Db	561834	CCGACAACAACATCGATTATCTGTGCGAAGCATCGTGAAGTACTGTAA	561785

RESULT 13
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429


```
Db 1684507 CAAAAATGTTGGWTCAGCACACCAACTTGGCCAAACCACAATGCGATTTTCAATGCTGTC 1684566
Qy 421 GGTCTGAAAGTTCGTGATACGCTTATTATGATCGCGAAATACACTCTTGTACTTCGAT 480
Db 1684567 GGTATGACCAATTCGTGAATATCGTATTATGATGCTGAACGCAAGCCCTTGATGGAA 1684626
Qy 481 GCACCTGATTAAACAGCCCTGAATGAAGCTCAGCGCTGGCGACCTAGTGCCTTCCATGGCTGC 540
Db 1684627 CACTTATTAGAGATTTTAAGCCAAGCAAGCAAGGCGATGTGTCTTTTACACGGTTGT 1684686
Qy 541 TGCATTAACCCAAACCGGTATCGACCCCTACGCTGGAAACAATGGCAAAACACTGGCACAACTC 600
Db 1684687 TGCCATAATCCGACTGCTGTTAGCCCTTACTCCAGAACAAATGGCAAGAAATTAGCGCACTT 1684746
Qy 601 TCGCTTAGAAAGGCTGGTTACCGCTGTTTGAAGCTTTCGCTTACCAGGCTTTTGGCCGCTGGT 660
Db 1684747 TCAGCTAAAAATGGTGTGGCTTGCCTACTCTTTGACTCTTATCAAGGTTTAGCCAAACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGGGCTTTTGGCGCTATGATATAAGAGCTGATTGTT 720
Db 1684807 TTAGATGAAGATGCTTATGTTTACGTGCTTTTGCAGCAAAACCAAGAAGAAATATTAGTG 1684866
Qy 721 GCCAGTTCTTACTCTATAAAACTTTGGCGCTGTACAAACAGAGCGTGTGGCGCTTGTACTCTG 780
Db 1684867 GCGAGTTCAATCTCGAAAAACTTTGGTTTATATAATGAACGTTGTGGTGCAATTTACCCCTT 1684926
Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTTAGCCAAATGAAGCGCGCATTCGC 840
Db 1684927 GTGCGCAAAATGCAGAAATTTGCATCAACCTCATTAACCCAGTAAATCAATTTATTCGC 1684986
Qy 841 GCTAACTACTCTAACCCACAGCACACGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
Db 1684987 ACCCTACTCTTAACCCAGCTTCTACGCGCGGGCGAGCCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATCGCTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
Db 1685047 GCTCAACTTCGCCAAGATGGGAATGAATTAACGAAATGCGTGAACGCATCAAAAA 1685106
Qy 961 ATGCGTCAGTTGTTGTCFAATAGCTGCAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1020
Db 1685107 ATGCGTCACATTATTCGTTCAAGTTATTAAGAAATATGGTCAGAAACAAGATTTTCAGCTTT 1685166
Qy 1021 ATCATCAACAGACGCGCATGTTCTCCTTCAGTGGCTGACAAAGAACAAAGTGCCTG 1080
Db 1685167 ATCATCAACAAACGCGTATGTTCTTTCAGTGGATTAACAGGGGAACAGTGGATCGT 1685226
Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGGCGCGGATG 1140
Db 1685227 TTAAGAAATGAATTTGCCATTTACGCTGTTCTGTTCTGTCGATCAACGTAGCTGGAATC 1685286
Qy 1141 ACACAGATTAACATGGCTCCGCTGTGCGAAGCGAATGTGGCAGTGCCTGTAA 1191
Db 1685287 ACAGAGATAATATTCGCTATCTATGTGAAAGTATCGTGAAGAGTACTTTAA 1685337

RESULT 14
US-10-158-865-1
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match      38.5%; Score 458.2; DB 8; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 1.1e-145;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCGGATTCGGGCGCTGGCGGATCTGTTT 60
Db 1684147 ATGTTTGACATATCAAGCGGACGAGCGATCCAATCTTAGCTTAGCGCAAGCATTT 1684206

Qy 61 CGTGGCGATGAACGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 1684207 AAATCCGAAACTCGCGAAATAAAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCA 1684266

Qy 121 GGCAGAACCCGGTACTGACCGAGGTGAAAAGGCTGAACAGTATCTGCTGAAAAGTAA 180
Db 1684267 GGCACACCCCAATTTATGACGCGGTAAAAGAGCCGCAAAACGATTTATTGATAAGGA 1684326

Qy 181 ACCACCAAAATTAACCTCGGCAATTCAGCGCATCCCTGAATTTGGTGTCTGCACTCAGAA 240
Db 1684327 AAAACCAAGATTTATCTGACTCTGATGATGATGCGGATTTACGCAACAAACAAAGCA 1684386

Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGACAAAGTGTCTCGCAGCGCACAGACT 300
Db 1684387 CTCCTTTTCGTTAAAGATTCTGAAGTCATCCAATCTAATCGAGCAAGACAGTACAAAGT 1684446

Qy 301 CCGGGGGGCACTGCGGCACACTACCGGTGGCTGCGGATTTCTCTGGCAAAAATACCAGGTT 360
Db 1684447 TTAGCGGGAACAGGTGCAATTACGCATTTGCGGCGAGAATTTATAACGCCCAAACTAAAGCA 1684506

Qy 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATAGAGCGTCTTTAACTCTGCA 420
Db 1684507 CAAAATGTTTGGWTTCAGCACCAACCTTGGCCAAACCAACATGCGATTTTCAATGCTGTC 1684566

Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACTTGTGCTTCGAT 480

```

```
Db 1684567 GGTATGACCATTCGTGAATATCGTTATTATGATGCTGAACGCAAAAGCCCTTGATGGGAA 1684626
Qy 481 GCAGTATTAACAGCCGTAATGAAGCTCAGCGTGGCGACGCTAGCTGCTTCCATGCTGC 540
Db 1684627 CACTTATTAGAAGATTTAAGCCAAGCAAGCGGATGTGTGCTTTTACACGGTGT 1684686
Qy 541 TGCATAAACCAACCGGTATCGACCCCTACGCTCGAACAATGGCAAAACACTGGGCACAACTC 600
Db 1684687 TGCATATATCCGACTGTGATTGACCCCTACTCCAGAACATGGCAAGAAATTAGCGGCACCT 1684746
Qy 601 TCGGTTGAGAAAGCGTGGTTACCGCTGTTTGACTTCGCTTACCAGGGTGTTCGCCGCTGGT 660
Db 1684747 TCAGTAAATAATGGTTGGTTGCCACTCTTTGACTTTTATCAAGGTTTAGCCACGGA 1684806
Qy 661 CTGGAAGAAATGCTGAAGACTCGCGGCTTTTCGGGCTATGATATAAGAGCTGATTGTT 720
Db 1684807 TTAGATGAAGATGCTTATGCTTTTACGCTGCTTTTGCAGCAAAACCAAGAAATTTATTAGTG 1684866
Qy 721 GCCAGTTCCCTACTCTAAATACTTTGGCGCTGTACAAAGAGCGTGTGGCGCTTGTACTCTG 780
Db 1684867 GCGAGTTCACTTCGAAAAAATCTTTGGTTTATATAATGAACGTGTGTGCAATTTACCCCT 1684926
Qy 781 GTTGCTGCCACAGTGAACCGTTGATCGGCGATTTCAGCCAAATGAAGCGGCGATTTCGC 840
Db 1684927 GTGCAGAAATGCAGAAATTCATCAACCTCATTAACCCAGTAAATCAATTTATTCG 1684986
Qy 841 GCTAACTACTCTAACCCACAGACACGCGCTTCTGTTTGTGCGACCATCTCTGAGCAAC 900
Db 1684987 ACCCTACTCTAACCCAGCTTCTCACGGGGGGCGACCGTAGCAACAGATTATAATGAT 1685046
Qy 901 GATCGGTAGTGTGGAATTCGGACACAGAGCTGACTGATATGCCAGCGTATTTACGCGT 960
Db 1685047 GCTCAACTCGCCAAAGAAATGGAAAAATGAATTAACGAAATGCGTGAACGCATCAAAAA 1685106
Qy 961 ATGCGTCAGTTGTCGTCATACGCTGCAGGAAAAAGCGGCAAAACCGCGACTTCAGCTTT 1020
Db 1685107 ATGGTCACATTATTCGTTTCAGTTATTAAGAATATGTTGCGAGAACAGATTTCAGCTTT 1685166
Qy 1021 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGTGACAAAGAACAAAGTGTGCGT 1080
Db 1685167 ATCATTAACAAACCGGTATGTTTCTTCAGTGGATTAAACAGGGGAACAAAGTGGATCGT 1685226
Qy 1081 CTGCGCAAGAGTTGCGGTATATGCGGTTGCTTCTGTCGCTTAATGTGGCGCGGATG 1140
Db 1685227 TTAATAAATGAATTTGCCATTACGCTGTTTCGTTCTGCTGATCAACGATGCTGGAATC 1685286
Qy 1141 ACACCATATACATGCTCGCTGTGCGAAGCGATTGTGGCAGTGTCTGTAA 1191
Db 1685287 ACAGAAGATATATTCGCTATCTATGTGAAGATATCGTGAAGTACTTTTAA 1685337

RESULT 15
US-10-981-687-1
; Sequence 1, Application US/10981687
; Publication No. US20050131222A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D12
; CURRENT APPLICATION NUMBER: US/10/981,687
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

Query Match      38.5%; Score 458.2; DB 9; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 1.1e-145;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTACCGCGGCTCTCGCGACCGGATTCGGCGCTTGGCGGCTGCGGCTGCTGTTT 60
Db 1684147 ATGTTTGAGAACATTACCGCGGCTCTCGCGACCGGATTCGGCGCTTGGCGGCTGCGGCTGCTGTTT 1684206

Qy 61 CGTGCCGATGAACGTCCCGGCAAAATTAACCTCGGGATTCGGTCTCTATAAGATGAGACG 120
Db 1684207 AATCCGAACTCGGAAATATAATTCATTTGGTATTGCGTTTATAAGATGCGCAA 1684266

Qy 121 GGCAAAACCCCGGTACTGACGCGTGCAAAAGGCTGAAACAGTATCTGCTCGAAATGAA 180
Db 1684267 GGCACACCCCAATTATGACGCGGTAAAGAGCCGAAACAGTATTTTGATAAGGAA 1684326

Qy 181 ACCACCAAAATTTACCTCGGGAATTGACGGGATCCCTGAAATTTGGTCTGCTGCACTCAGAA 240
Db 1684327 AAAACCAAGAAATTTATCTGACTATCGATGTTTGGGATTTATACGAAACAAACAAAGCA 1684386

Qy 241 CTGCTGTTTGGTAAAGGTAGCGGCTGATCAATGACAAAGCTGCTCCGACGGCACAGACT 300
Db 1684387 CTCCTTTTCGGTAAAGATTCGAAATCTCAATCTAATCGAGCAAGACAGTACAAAGT 1684446

Qy 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTCTGGCAAAATAATACCAGCGTT 360
Db 1684447 TTAGCGGAACAGGTGCATTTACGCATTTGCGGAGAAATTTATTAACCCCAACTAAAGCA 1684506

Qy 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGCGGAAACCATAGAGCGCTCTTTAACTCTGCA 420
Db 1684507 CAAAATGTTTGGWTCAGCACACCAACTTGGCCAAACCAACATCGAATGCTGCTGCTGTC 1684566

Qy 421 GGTCTGGAAGTTCTGTGAATAGCTTTATATGATCGGAAATACACCTCTTGACTTCGAT 480
Db 1684567 GGTATGACCAATTCGTGAATATCGTTATTTATGATGCTGAACGCAAGCCCTTGATGGAA 1684626

Qy 481 GCACGTATTAAACAGCCCTGGAATGAAGCTCAGCTGCGGACGATGCTGTTTCCATGGCTGC 540
```

```

Db 1684627 CACTTATTAGAGATTTAAGCCAAAGCAAGCGGATGTGGTGTCTTTTACACGGTTGT 1684686
Qy 541 TGCCTAAACCCACCGGTATCGACCTTACGCTGGAAACATGGCAACACTGGGCACAACTC 600
Db 1684687 TGCCTAAATCCGACTGGTATTGACCTTACTCCAGAAACAATGGCAAGAAATTAGCCGACAT 1684746
Qy 601 TCCGTTGAGAAAGGCTGGTTTACCGCTGTTTGAATCTTGGCTTACCAGGCTTTTGCCCGTGGT 660
Db 1684747 TCAGCTAAATAATGGTTGGTTGCCACTCTTTGACTTTTGCCTTATCAAGTTTAGCCAACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTCCGCGCTATGCAATAAGAGCTGATTTGT 720
Db 1684807 TTAGATGAAGATGCTTATGCTTTTACGTGCTTTTGCAGCAAAACCAAGAATATTAGTG 1684866
Qy 721 GCCAGTTCCTACTCTAAATACTTTGGCTGTACAAAGAGCGTGTGGCGCTTGTTACTCTG 780
Db 1684867 GCGAGTTTCATCTCGAAAAAATTTGGTTTATATAATGAACGCTTGGTGCAITTTACCCCT 1684926
Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTTAGCCAAATGAAGCGGCGATTTCGC 840
Db 1684927 GTGGCAAAATTCGAGAAATTTGATCACTCACTTAAACCAAGTAAATCAATTATTCGC 1684986
Qy 841 GCTAACTACTCTAACCCACAGCACACGCGCTTCTGTGTTGTCACCAATCCTTGAGCAAC 900
Db 1684987 ACCCTATACTCTAACCCAGCTTCTCACGGCGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 901 GATGCGTTACGTGGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
Db 1685047 GCTCAACTTCGCCAAGAATGGGAAAATGAATTAACCTGAATGCGTGAAACGCATCAAAAA 1685106
Qy 961 ATGCGTCAGTTGTTTCGTCATACGCTGCAGGAAAAAGGCGCAAAACCGGACTTTCAGCTTT 1020
Db 1685107 ATGCGTCACCTATTTCGTTTCAGTTATTAAAGAATATGGTGCAAGCAAGATTTTCAGCTTT 1685166
Qy 1021 ATCATCAAAACAGAACGGCATGTTCTCCTTCAGTGGCTGCAAAAAAGAACAAAGTCTGCGT 1080
Db 1685167 ATCATTGAAACAAACCGTATGTTTCTTTTCAGTGGATTAAACAGGGGAACAAGTGGATCGT 1685226
Qy 1081 CTGGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGGTCGCGTAAATGTGGCGGGATG 1140
Db 1685227 TTAATAAATGAATTTGCCATTTACGCTGTTTCGTTCTGTCGTATCAACGCTAGCTGGAATC 1685286
Qy 1141 ACACCAAGATAACATGGCTCGCTGTGGAAGCGGATTTGTCAGTGTGCTGTAA 1191
Db 1685287 ACAGAAGATAATATTCGCTATCTATGTGAAAGTATCGTGAAGTACTTTAA 1685337

```

Search completed: March 15, 2006, 15:45:16
Job time : 2988 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 03:03:01 ; Search time 1180 Seconds
(without alignments)
2236.626 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045
Sequence: 1 MFENITAPADPILGLDLF.....VAGTPDNPALCEAIVAVL 396

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10673786/runat_14032006_135608_22579/app_query.fasta_1
-DB=N Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10673786 @CGN 1 1 1929 -runat_14032006_135608_22579 -NCPUs=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DEPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1191	9 ADB83284	Ad83284 Escherich
2	2045	100.0	1191	13 ADS46042	Ads46042 Bacterial
3	2045	100.0	1191	14 ADW95403	Adw95403 Nucleotid
4	2045	100.0	1331	2 AAV40259	Aav40259 Escherich

5	2045	100.0	3659	1	AAN71109	Aan71109 pheA aroF	
6	2045	100.0	14759	4	AAS46273	Aas46273 DNA encoded	
c	7	2039	99.7	1	AAN71108	Aan71108 Optimized	
8	1931	94.4	1209	11	ACH99107	Ach99107 Klebsiell	
9	1682	82.2	1191	10	ACF68649	Acf68649 Photorhab	
10	1682	82.2	110000	10	ACF67367_14	Continuation (15 o	
11	1682	82.2	249878	10	ACF65381	Acf65381 Photorhab	
12	1654	80.9	1188	13	ADT46374	Adt46374 Bacterial	
13	1343	65.7	7977	13	ADT05500	Adt05500 Haemophil	
c	14	1343	65.7	349980	13	ADT05649	Adt05649 Haemophil
15	1325	64.8	110000	2	AAT42063_16	Continuation (17 o	
16	1318	64.4	966	13	ADS45488	Ads45488 Bacterial	
17	1317	64.4	1269	10	ABZ40973	Abz40973 N. gonorr	
18	1311	64.1	1191	6	AAS97289	Aas97289 Neisseria	
19	1310	64.1	17381	3	AAA81493	Aaa81493 N. mening	
c	20	1310	64.1	110000	3	AAA81490_05	Continuation (6 of
c	21	1310	64.1	349980	3	AAF21607_	AAF21607 Neisseria
22	1005.5	49.2	1185	13	ADS59837	Ads59837 Bacterial	
23	996	48.7	1188	13	ADS55510	Ads55510 Bacterial	
24	996	48.7	1194	13	ADS49995	Ads49995 Bacterial	
25	982	48.0	1191	13	ADS56437	Ads56437 Bacterial	
26	982	48.0	1242	11	ABD13824	Abd13824 Pseudomon	
c	27	982	48.0	1329	11	ABD13558	Abd13558 Pseudomon
c	28	982	48.0	24417	2	AAT97221	Aat97221 Pseudomon
29	975	47.7	1191	13	ADS63841	Ads63841 Bacterial	
30	975	47.7	1191	13	ADS64221	Ads64221 Bacterial	
31	975	47.7	1194	13	ADS63473	Ads63473 Bacterial	
c	32	955	46.7	579	11	ACH99159	Ach99159 Klebsiell
33	947	46.3	1194	13	ADS62038	Ads62038 Bacterial	
34	924	45.2	1185	13	ADS57471	Ads57471 Bacterial	
35	917	44.8	1185	13	ADS57235	Ads57235 Bacterial	
36	917	44.8	1481	13	ADX64272	Adx64272 Plant full	
37	914	44.7	1489	13	ADX45580	Adx45580 Plant full	
38	913	44.6	1456	13	ADX47271	Adx47271 Plant full	
39	911.5	44.6	1185	13	ADS61736	Ads61736 Bacterial	
40	911	44.5	1508	13	ADX60438	Adx60438 Plant full	
c	41	909.5	44.5	1509	11	ABD14364	Abd14364 Pseudomon
42	909.5	44.5	1632	11	ABD13903	Abd13903 Pseudomon	
43	909	44.4	1203	13	ADT42736	Adt42736 Bacterial	
44	908	44.4	702	10	ADF03520	Adf03520 Bacterial	
45	893	43.7	1188	13	ADS56672	Ads56672 Bacterial	

ALIGNMENTS

RESULT 1
ADB83284
ID ADB83284 standard; DNA; 1191 BP.
XX
AC ADB83284;
XX
DT 04-DEC-2003 (first entry)
XX
DE Escherichia coli aspartate aminotransferase gene.
XX
KW ds; gene; aspartate aminotransferase; threonine; fermentation.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..1191
FT /tag= a
FT /product= "aspartate aminotransferase"
XX
PN WO20003072786-A1.
XX
PD 04-SEP-2003.
XX
PF 25-FEB-2003; 2003WO-JP002067.
XX
PR 27-FEB-2002; 2002RU-00104983.
XX
PA (AJIN) AJINOMOTO CO INC.

XX Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;
 XX WPI; 2003-721777/68.
 XX P-PSDB; ADB83285.
 XX Industrial production of L-threonine by fermentation using *Escherichia*
 PT modified to enhance aspartate aminotransferase activity, with improved
 PT productivity.
 XX Claim 7; Page 18-20; 26pp; Japanese.
 XX The invention relates to a bacterium belonging to the genus *Escherichia*
 CC which is modified to enhance aspartate aminotransferase activity for the
 CC production of L-threonine. The bacterium is used for the industrial
 CC production of L-threonine by fermentation. This sequence corresponds to
 CC the *E. coli* aspartate aminotransferase gene.
 XX
 SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8,798-203 Length: 1191
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x ADB83284 (1-1191)

QY 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 DB 1 ATGTTTGAGAACATTACCGCGCTCTCGCCGACCGGATCTGGGCTGGCGGATCTGTTT 60
 QY 21 ArgAlaAspGluAArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 DB 61 CGTCCGATGAACGTCGCGGCAAAATTAACTCGGGATTGGTGCTATAAAGATGAGACG 120
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu 60
 DB 121 GCGAAACCCCGGTACTACGACGAGGTGAAAGGCTGAACAGATATCTGCTCGAAATGAA 180
 QY 61 ThrThrLysAsnTyrLeuGlyLysAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 181 ACCACCAAAATTAACCTCGCATTCGCGCATCCCTGAATTTGGTCGTCGACACAGGAA 240
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 241 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCACGCGACAGACT 300
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 301 CCGGGGGGCACTGGCGCACTACGGGTGGCTGCCGATTTCCCTGGCAAAAATACACAGCTT 360
 QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 DB 361 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGCAACCAATAAGAGCGTCTTTAACTCTGCA 420
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAAsnHisThrLeuAspPheAsp 160
 DB 421 GGTCTGGAAGTTCTGGTAATAGCGCTATTATGATCGGAAATATCACTCTTTGACTTCGAT 480
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 DB 481 GCATGATTAAACGCTGAATGAAGCTCAGCTGGCGACGTAGTGCTGTTCCATGGCTGC 540
 QY 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 DB 541 TGCCATAACCCCAACCGGTATCGACCTACGCTAGCTGGAAACAATGGCAACACACCTGCAACTC 600
 QY 201 SerValGluLysGlyTyrProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 DB 601 TCCGTTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACCAGGGTTTGGCCCGTGT 660

QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 DB 661 CTGGAAGAAGATGCTGAAGGACTGCGGCTTTTCGGCTATGCATAAAGAGCTGATTGTT 720
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 DB 721 GCCAGTTTCTTACTCTCTAAAAACTTTGGCTGTGTACAAACGAGCGGTGTGGCGCTTGACTCTG 780
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCCAATGAAAGCGCGCATTCGC 840
 QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 DB 841 GCTAACTACTCTAAACCCACGACACGCGGCTTCTGTGTGTGCCACCATCTCTGAGCAAC 900
 QY 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 901 GATCGGTTTACGTGCGATTTCGGGAACAAGAGCTGACTGATATGCCGCGGATTCAGCGT 960
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPhe 340
 DB 961 ATGGCTCAGTTGTTCGTCATACGCTCAGGAAAGGCGCAACCCGACTTCAGCTTT 1020
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 1021 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT 1080
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 DB 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTCCTTCGTCGCGTAAATGTGGCCGGGATG 1140
 QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 1141 ACACAGATAACATGGCTCCGCTGTCGCAAGCGATTGTGGCAGTGCTG 1188

RESULT 2
 ADS46042
 ID ADS46042 standard; cDNA; 1191 BP.
 XX
 AC ADS46042;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #785.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 OS
 PN US20032333675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PR (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 FI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX

DR WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 24472; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at: seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,79e-203 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADS46042 (1-1191)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 1 ATGTTTGAGAACATTACCGCGCTCTGCGGACCGGATTCCTGGGCTGGCCGATCTGTTT 60
QY 21 ATGAlaAspGluArgProGlyIleAsnLeuGlyIleGlyValTyrlsAspGluThr 40
DB 61 CTGGCGGATGAACGTCCTGGGCAAAATTAACCTCGGATTTGGTGTCTATTAAGATGAGAG 120
QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyLzLeuLeuGluAsnGlu 60
DB 121 GCGAAACCCCGGTACTACGAGCGTGAAGAGGCTGAACATCTCTGCTCGAATGNA 180
QY 61 ThrThrLysAsnTyrlsLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 181 ACCACCAAAATTAACCTCGGATTCGCGCATCCCTGAATTTGGTTCGCTGCACTCAGGAA 240
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
DB 241 CTGCTGTTTGGTAAAGGTAGCCCTGTATCAATGACAAACGTCGTCGACGCGCACAGACT 300
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 301 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGATTTCTGGCAAAATACACGAGTT 360
QY 121 LysArgValTrrpValSerAsnProSerTrrpProAsnHisLysSerValPheAsnSerAla 140
DB 361 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGGACCAATACAGAGCGTCTTTAACTCTGCA 420

QY 141 GlyLeuGluValArgGluTyrlsAlaTyrlsAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 421 GGTCTGGAGATTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTGACTTCGAT 480
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 481 GCACTGATTAAACAGCGCTGAATGAAGCTCAGGCTGCGAGCTAGTGTGTTCATGCTGCTC 540
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
DB 541 TGCCATTAACCCCAACCGGTATCGACCTAGCTGGAAACAATGGCAACACTGGCACAACTC 600
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrlsGlnGlyPheAlaArgGly 220
DB 601 TCCGTTTGAGAAAGGCTGGTTACCGCTGTTTACCTTCGCTTACCAGGGTTTGGCCGCTGT 660
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
DB 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGGCTATGCATAAAGAGCTGATTGT 720
QY 241 AlaSerSerTyrlsSerLysAsnPheGlyLeuTyrlsAsnGluArgValGlyAlaCysThrLeu 260
DB 721 GCAGTCTCTACTCTTAAACACTTTTGGCCTGTACAACGAGCGTGTGGCCGCTTGACTCTG 780
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
DB 781 GTTGTGCGGACAGTGAACCGTTCATCGCGCATTCAGCCAATGAAGCGGCGATTCGC 840
QY 281 AlaAsnTyrlsSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
DB 841 GCTAACTACTCTTAACCCACAGCACACGCGCTTCTGTGTGGCCACCATCTCTGAGCAAC 900
QY 301 AspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
DB 901 GATGCTTACGTCGCGATTTGGGAAACAAGAGCTGACTGATATGCGCAGCGTATTTCAGCG 960
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAsnArgAspPheSerPhe 340
DB 961 ATGCGTCACTGTTCTCAATACGCTGAGGAAAGGCGCAACCGCGACTTCAGCTTT 1020
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
DB 1021 ATCATCAAAACAGAACGCGATGTTCTCTTCACTGCGCTGACAAAGAACCAAGTGTGCTG 1080
QY 361 LeuArgGluGluPheGlyValTyrlsAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 1081 CTGCGCGAAGAGTTTGGCGTATATCGGCTATATCGGTTGCTTCTGTCGCTAAATGTGCGCGGATG 1140
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
DB 1141 ACACCAAGATACATGCTCCGCTGTCGAGCGATTTGGCAGCTGCTGCTG 1188
RESULT 3
ID ADW95403 standard; DNA; 1191 BP.
XX
AC ADW95403;
DT 07-APR-2005 (first entry)
XX
DE Nucleotide sequence of an Escherichia coli polypeptide.
XX transgenic plant; glutamic acid dehydrogenase; GDH; ECASPC;
KW 2-oxo glutaric acid; agriculture; gene; ds.
XX Escherichia coli.
OS
FH Key Location/Qualifiers
FT CDS 1..1191
XX /*tag= a
PN WO2005006847-A1.
XX

PD 27-JAN-2005.
 XX 15-JUL-2004; 2004WO-JP010451.
 XX PF
 XX PR
 XX 17-JUL-2003; 2003JP-00198559.
 XX (AJIN) AJINOMOTO CO INC.
 XX PA
 XX KI
 XX Kisaaka H, Miwa T, Akiyama A;
 XX WPI; 2005-132242/14.
 XX DR P-PSDB; ADW95404.
 XX PT
 XX Producing plant having improved growth and yield under cultivation
 PT conditions with decrease in nitrogen, involves expressing transduced
 PT glutamic acid dehydrogenase gene in plant so as to increase their 2-oxo
 PT glutaric acid contents.
 XX
 XX Disclosure; SEQ ID NO 19; 94pp; Japanese.
 XX
 XX The specification describes a method of producing a plant having improved
 CC growth and yield under cultivation conditions with a decrease in
 CC nitrogen. The method involves transducing glutamic acid dehydrogenase
 CC (GDH) or ECASPC genes into a plant and expressing the gene so as to
 CC increase 2-oxo glutaric acid content of the plant, or applying proline to
 CC the foliage of the plant so as to increase the 2-oxo glutaric acid
 CC content of the plant. The method is useful for producing a plant having
 CC improved growth and yield under cultivation conditions with a decrease in
 CC nitrogen. The present sequence encodes a polypeptide, used in the course
 CC of the invention.
 XX
 XX Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 8,79e-203 Length: 1191
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0
 US-10-673-786A-2 (1-396) x ADW95403 (1-1191)
 Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 1 ATGTTTGAGAACATTACCGCGCTCTCCGACCGGATTCCTGGCGCTGGCGGATCTGTTT 60
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 61 CGTCCGATGAACGTCCTCCGCAAAATTAACCTCGGGATTTGGTGTCTATAAAGATGAGACG 120
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 121 GGCAAAACCCCGGTACTGACGAGCGTGAAAGGCTGACAGATATCTGCTCGAAATGAA 180
 Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 181 ACCACCAAAATTAACCTCGCATTCGCGCATCCCTGAATTTGGTGTCTGCTGACCTCAGGAA 240
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 300
 Qy 101 ProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 301 CCGGGGGGACCTGGCGCACTACGCGTGCTGCCGATTTCTGGCAAAAATACACGCGTT 360
 Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 361 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGCAACCATTAAGAGCGTCTTTAACTCTGCA 420
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160

Db 421 GGTCGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTTGACTTCGAT 480
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTCTTTCATGGCTGC 540
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 541 TGGCATAAACCCACCGGTATCGACCTTACCGCTGACCAATGGCAACACTGGGCAACAATC 600
 Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 601 TCCGTTGAGAAAGGCTGTTTACCGCTGTTGACTTTCGCTTACCAAGGTTTTTGGCCGTGT 660
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 720
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGTTTCTTACTCTAAAAACTTTGGCCTGTACACGAGCGTGTGGCGCTTTGACTCTG 780
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGTGCGCCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGATTCGC 840
 Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 841 GCTAACTACTCTAACCCACGACGACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgGlnArg 320
 Db 901 GATCGGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCCCGACGCTATTCAGCGT 960
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 Db 961 ATGGCTCAGTTTGTCTCAATACGCTCAGGAAAGGCGCAACCGGACTTTCAGCTTT 1020
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATCATCAACAGAACGCGATGTTCTCTTCACTGCGCTGACAAAGAACCAAGTCTGCGT 1080
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTGCTTCTGCGTAAATGTGCGCGGATG 1140
 Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACCAATTAACATGGTCCGCTGTCGAGCGATTTGGCAGTGTCTG 1188
 RESULT 4
 AAV40259
 ID AAV40259 standard; DNA; 1331 BP.
 XX
 AC AAV40259;
 XX
 DT 13-OCT-1998 (first entry)
 XX
 DE Escherichia coli aspC gene.
 XX
 XX Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
 KW aspartokinase; feedback inhibition; dihydropicolinate reductase;
 KW diaminopimelate decarboxylase; aspartate aminotransferase; ds.
 XX
 OS Escherichia coli.
 XX
 XX Location/Qualifiers
 FH 10. .1197
 CDS /tag= a
 FT /product= "aspC"
 XX
 PN EP854189-A2.
 XX

PD 22-JUL-1998.
 XX
 PF 05-DEC-1997; 97EP-00121443.
 XX
 PR 05-DEC-1996; 96JP-00325659.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 PI Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;
 XX
 DR WPI; 1998-379060/33.
 DR P-PSDB; AAW69553.
 XX
 PT Recombinant DNA autonomously replicable in coryneform bacteria - used to
 PT produce L-lysine, codes for e.g. aspartokinase, di:hydronicotinase
 PT reductase and synthase and di:amino-pimelate decarboxylase.
 XX
 XX Example 5; Page 37-38; 59pp; English.
 XX
 CC The present invention describes a recombinant DNA autonomously replicable
 CC in cells of coryneform bacteria (CB), comprising a DNA sequence coding
 CC for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
 CC threonine is desensitized, a DNA sequence coding for a
 CC dihydronicotinase reductase (DHPR), a DNA sequence coding for
 CC dihydronicotinase synthase (DHPS), a DNA sequence coding for
 CC diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
 CC aspartate aminotransferase (AAT). The present sequence encodes aspC from
 CC Escherichia coli. The DNA and related products from the present
 CC invention, can be used for improving L-lysine productivity by CB. The L-
 CC lysine produced can be used as a fodder additive
 XX
 SQ Sequence 1331 BP; 330 A; 340 C; 350 G; 311 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: 1,03e-202 Length: 1331
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-673-786A-2 (1-396) x AAV40259 (1-1331)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 10 ATGTTTGAGAACATTACCGCGCTCTGCGCCGACCGGATTCGGCGCTGGCCGATCTGTTT 69
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyLysIleValTyrLysAspGluThr 40
 Db 70 CGTCCGATGAACGTCCCGCAAAATTAACCTCGGATTTGTTGCTATAAAGATGAGACG 129
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuGluAsnGlu 60
 Db 130 GGCACAAACCCCGGTACTGACACGCGTGAAGGCTGAACAGATATCTGCTCGAAATGAA 189
 Qy 61 ThrThrLysAsnTyrLeuGlyLysIleAspGlyLysIleProGluPheGlyArgCysThrGlnGlu 80
 Db 190 ACCACCAAAATTAATTCCTCGGATTTGACGCGATCTCTGAAATTTGGTGGCTGACCTCAGGAA 249
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 250 CTGCTGTTTGTAAAGGTAGCGCTGATCAATGACAAACGCTGCGACGCGACAGACT 309
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 310 CCGGGGGGCACTGGCGCACTACGGGTGGCTGCCGATTTCTGGCAAAAAATACACGCGTT 369
 Qy 121 LysArgValTyrValSerAsnProSerTrpProHnIleLysSerValPheAsnSerAla 140
 Db 370 AAGCGTGTGGGTGAGCAACCCAGCTGGCGGACCAATAGAGCGCTTTTAATCTGTGCA 429
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAspPheAsp 160

Db 430 GGTCTGGAAGTTCGTGAATAGCTTATTATGATCGCGAAATACACTCTTGTGACTTCGAT 489
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 490 GCACTGATTAAACAGCTGAATGAAGCTCAGCTGCGACGCTAGTGTGTTCCATGGCTGC 549
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 550 TGCCATAACCAACCGGTATCGACCTACGCTGGAACAATGCAAACTGCGCAACTC 609
 Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 610 TCCGTTTGAGAAAGCTGTTTACCGCTGTTTACCTTCGCTTACCAAGGTTTTCGCCGCTGT 669
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 Db 670 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGGGCTATGCATAAAGAGCTGATTCTT 729
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 730 GCCAGTTCTTACTCTAAAACTTTGGCTGTGTACACAGCGTGTGGCGCTTGTACTCTG 789
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 790 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGC 849
 Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 850 GCTAACTACTTAAACCCACAGCACAGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 909
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 910 GATCGGTTACGTGCGATTTGGGAACAAGAGCTGATGATGCGCCAGCGTATTTCAGCGT 969
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGluGlyLysGlyAlaAsnArgAspPheSerPhe 340
 Db 970 ATGCGTCAGTGTGTTGCTCAATACCTGCGAGAAAGCGCAACCCGCGACTTCAGCTTT 1029
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1030 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTCTGCGT 1089
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1090 CTGCGCGCAAGAGTTTGGCTGATATGCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1149
 Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1150 ACACAGATAACATGGCTCGCTGCGAAGCGATTGTGGCAGTGCTG 1197

RESULT 5
 AAN71109
 ID AAN71109 standard; DNA; 3659 BP.
 XX
 AC AAN71109;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE pheA aroF aspC operon in plasmid pME219.
 XX
 KW alpha-amylase; feedback inhibition; amino acid synthesis;
 KW composite plasmid; ss.
 XX
 OS Bacillus licheniformis.
 XX
 PH Key Location/Qualifiers
 CDS 69..1241
 FT /*tag= a
 FT 1259..2329
 FT /*tag= b
 FT 2344..3534
 FT /*tag= c
 XX

```

PN W08700202-A.
XX
PD 15-JAN-1987.
XX
PF 24-JUN-1986; 86WO-US001353.
XX
PR 24-JUN-1985; 85US-00747732.
XX
PA (NUTR-) NUTRASWEET CO.
XX
PI Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
XX WPI; 1987-021998/03.
XX DR P-PSDB; AAP70752, AAP71677, AAP71678.
XX
XX Composite plasmids contg. multiple genes in transcriptional units -
PT useful for prodn. of amino acid(s), esp. L-phenylalanine and L-tyrosine.
XX
XX Disclosure; Page 38; 57pp; English.
XX
XX This sequence may be inserted into a composite plasmid and used for the
CC production of amino acids. See also AAN71053-55, AAN71107, AAN71109,
CC AAN71111 and AAP70696-97 and AAP70750, AAP70752 and AAP70754
XX
XX Sequence 3659 BP; 936 A; 911 C; 947 G; 865 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred No.: 4.19e-202 Length: 3659
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x AAN71109 (1-3659)
Qy 1 MetPheGluAenIleThrAlaAProAlaAAspProIleLeuGlyLeuAlaAAspLeuPhe 20
XX
Db 2344 ATGTTTGAGAACATTACCGCGCTCTGCCGACCGGATTCCTGGCGCTGGCGGATCTGTTT 2403
Qy 21 ArgAlaAspGluArgProGlyIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 2404 CGTCCGATGAACGTCCCGGAAAAATTAACCTCGGGATTGGTGCTATAAAGATGAGACG 2463
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu 60
Db 2464 GGCAAAACCCCGTACTGACCGCGCTGACCGGATCCCTGAATTTGGTCTGCACTCAGGAA 2523
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 2524 ACCACCAAAATTTACCTCGGATTCGCGGATCCCTGAATTTGGTCTGCACTCAGGAA 2583
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 2584 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATACAAACGCTGCTCGACGCGACAGACT 2643
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAAspPheLeuAlaLysAAsnThrSerVal 120
Db 2644 CCGGGGGGCACTGGCGGACATACGGCTGCTGCCGATTTCTTGGCAAAAAATACCGCGTT 2703
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAAsnSerAla 140
Db 2704 AAGCGTGTGGTGAGCAACCCCAAGCTGGCGGACCATTAAGAGCGCTTTTAACTCTGCA 2763
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAAspPheAsp 160
Db 2764 GGTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAAAATCACACTCTTGACTTCGAT 2823
Qy 161 AlaLeuIleAenSerLeuAAsnGluAlaGlnAlaGlyAAspValValLeuPheHisGlyCys 180
Db 2824 GCACGTGATTAAACCCCTGAATGAAGCTCAGGCTGGCGACGATGCTGCTTCATGGCTGC 2883
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200

```

```

Db 2884 TGCATAAACCCCGGTATCGACCTTACGACCAATGGGCAACACTGGGCAACTC 2943
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 2944 TCCGTTGAGAAAGGCTGTTACCGCTGTTGACTTCGCTTACCAGGGTTTGGCCGTTGT 3003
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleVal 240
Db 3004 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTCGCGCTATGCATAAAGAGCTGATTGTT 3063
Qy 241 AlaSerSerTyrSerLysAAsnPheGlyLeuTyrAAsnGluArgValGlyAlaCysThrLeu 260
Db 3064 GCCAGTTCTTACTCTAAAAACTTTGGCCTGTACAACGAGCGTGTGGCGCTTGACTCTG 3123
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 3124 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGGCGGATTCGC 3183
Qy 281 AlaAsnTyrSerAAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAAsn 300
Db 3184 GCTAACTACTCTTAACCCACGACACGCGGCTTCTGTTGTTGCCACCATCTTGAGCAAC 3243
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 3244 GATCGCTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCACGCT 3303
Qy 321 MetArgGlnLeuPheValAAsnThrLeuGlnGluLysGlyAlaAAsnArgAAspPheSerPhe 340
Db 3304 ATCGCTCAGTTGTTGTTCAATACGCTGCAGGAAAGGCGCAAAACCGCGACTTCAGCTTT 3363
Qy 341 IleIleLysGlnAAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 3364 ATCATCAACACAGAACGCGATGTTCTCTTCACTGGCTGACAAAGAAACAAGTCTCGCT 3423
Qy 361 LeuArgGluPheGlyValTyrAlaValAlaSerGlyArgValAAsnValAlaGlyMet 380
Db 3424 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGGTCCGCTAAATGTGGCGGGATG 3483
Qy 381 ThrProAAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 3484 ACACAGATAACATGGCTCCGCTGTCGGAAGCGATTGTGCGCAGTGCTG 3531

RESULT 6
AAS46273/c
ID AAS46273 standard; DNA; 14759 BP.
XX
AC AAS46273;
XX
DT 18-DEC-2001 (first entry)
XX
DE DNA encoding novel mar regulated protein (NIMR) #42.
XX
KW mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.
XX
OS Escherichia coli.
XX
PN WO200170776-A2.
XX
PD 27-SEP-2001.
XX
PF 08-MAR-2001; 2001WO-US007478.
XX
PR 10-MAR-2000; 2000US-0188362P.
XX
PA (TUFT ) TUFTS COLLEGE.
XX
PI Levy SB, Barbosa TM, Alekshun MN;
XX
DR WPI; 2001-602769/68.
XX
DR P-PSDB; AAU29374.
XX
PT Identifying compounds that modulate a newly identified mar regulated

```

polypeptide activity, useful as antimicrobial compounds, involves contacting the polypeptide with a test compound.

Disclosure; Page 477-485; 526pp; English.

The invention relates to a method of identifying compounds that modulate a newly identified mar regulated (NMR) polypeptide activity. The method comprises contacting an NMR polypeptide with a test compound under interaction conditions, determining the ability of the compound to modulate the activity or expression of the polypeptide, and selecting the modulators. NMR nucleic acids and polypeptides are used in the treatment of microbial infections, and in screening for modulators of NMR expression and activity. These modulators can be used to reduce the infectivity of a microbe on a surface, and the virulence of a microbe in a subject suffering from an infection. AAS46273-AAS46278 represent *Escherichia coli* NMR coding sequences of the invention

Sequence 14759 BP; 3703 A; 3840 C; 3713 G; 3503 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.92e-201 Length: 14759
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-673-786A-2 (1-396) x AAS46273 (1-14759)

Qy	1	MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	2755	ATGTTTGAACAATTTACCGCGCTCTCGCCAGCCGATTCTGGGCTGGCGATCTGTTT	2696
Qy	21	ArgAlaAspGluAArgProGlyIleAsnLeuGlyIleGlyValTyrLysAspGluThr	40
Db	2695	CGTCCGATGAACGTCGCCGGAATAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG	2636
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu	60
Db	2635	GGCAAAACCCGGTACTACCCAGCGTGAATAAGGCTGAACAGTATCTCTCGAATAAGAA	2576
Qy	61	ThrThrLysAsnTyrLeuGlyLysLeuGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	2575	ACCACCAAAATTTACCTCGCATTCGACGCGATCCCTGAATTTGGTCGCTGCACTCAGAA	2516
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	2515	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT	2456
Qy	101	ProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	2455	CCGGGGGCACTGGCGCACTACCGCTGCTGCCGATTTCTTGGCNAANAATACCCAGCGTT	2396
Qy	121	LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
Db	2395	AAGCGTGTGGTGAGCAACCCAGCTGGCGGAACCAATAAGAGCGTCTTTAACTCTGCA	2336
Qy	141	GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	2335	GGTCTGGAAGTTCTGGAATACGCTTATTATGATGCGGAATAATCACACTCTTGTACTTCGAT	2276
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	2275	GCATGTATTACAGCCTGAATGAAGCTCAGGCTGGCGACGAGTAGTCTGTTCCATGGCTGC	2216
Qy	181	CysHisAsnProThrGlyLysAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
Db	2215	TGCCATAACCAACCGGTATCGACCTAGCTGGAGCAATGGAACAACTGGCACAACCTC	2156
Qy	201	SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	2155	TCCGTTGAGAAGGCTGTTACCGCTGTTTACCTTCGCTTACCGAGGTTTTCGCCGCTGGT	2096

Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
Db	2095	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTCGGCGTATGCATAAAGAGCTGATGTT	2036
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValaCysThrLeu	260
Db	2035	CCAGTTTCTTACTCTTAAAACTTTTGGCTGTACAAAGAGCGTGTGGCGCTTGTACTCTG	1976
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
Db	1975	GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGATTGCG	1916
Qy	281	AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn	300
Db	1915	GCTAACTACTCTAACCCACACGACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	1856
Qy	301	AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
Db	1855	GATCGTTTACGTGCGATTGGGAACAAGAGCTGATATGCGCCAGCGATTTCAGCGT	1796
Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValaAsnArgAspPheSerPhe	340
Db	1795	ATGCGTCACTGTTGCTCAATACGCTGCAGGAAAAGGCGCAACCCGCGACTTCAGCTTT	1736
Qy	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	1735	ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGACAAAAGAACAGTGTGCTGCT	1676
Qy	361	LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet	380
Db	1675	CTGGCGGAAGAGTTGGCGTATATGCGGTTCCTTCGTGCGTAAATGTGGCCGGGATG	1616
Qy	381	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	1615	ACACAGATAACATGCTCCGCTGTGCGAAGCGATTGTGGCAGTGTCTG	1568

RESULT 7

AAAN71108	ID	AAAN71108 standard; DNA; 1293 BP.
XX	AC	AAAN71108;
XX	DT	01-JAN-1980 (first entry)
XX	DE	Optimized <i>Escherichia coli</i> aspC gene.
XX	KW	aspC gene; feedback inhibition; amino acid synthesis; composite plasmid;
XX	XX	ss.
XX	OS	<i>Escherichia coli</i> .
XX	PN	W08700202-A.
XX	PD	15-JAN-1987.
XX	PF	24-JUN-1986; 86WO-US001353.
XX	PR	24-JUN-1985; 85US-00747732.
XX	PA	(NUTR-) NUTRASWEET CO.
XX	PI	Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
XX	DR	WPI; 1987-021998/03.
XX	DR	P-ESDB; AAP70751.
XX	PT	Composite plasmids contg. multiple genes in transcriptional units -
XX	XX	useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.
XX	PS	Disclosure; Page 25; 57pp; English.
XX	CC	This sequence may be inserted into a composite plasmid and used for the
CC	CC	production of amino acids. See also AAAN71053-55, AAAN71107, AAAN71109-11

```
CC and AAP70696-97 and AAP70750, AAP70752-54
XX Sequence 1293 BP; 322 A; 331 C; 343 G; 297 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 4,15e-202 Length: 1293
Score: 2039.00 Matches: 395
Percent Similarity: 99.7% Conservations: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x AAN71108 (1-1293)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 12 ATGTTTGAGAACATTACCGCGCTCCTGCCACCGATTCTGGCGCTGGCGATCTGTTT 71
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 72 CGTCCGATGAACGTCCCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 131
Qy 41 GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 132 GGCAAAACCCCGGTACTGACCGCGTGAAGAGCTGAACAGTATCTGCTCGAAATGAA 191
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 192 ACCACCAAAATTAACCTCGGATTCAGCGCATCCTGAAATTTGGTGTGCTGCACACAGAA 251
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 252 CTGCTGTTTGTAAAGTAGCGCTGATCAATGACAAACGTGCTCGCAGCGCACAGACT 311
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 312 CCGGGGGGCGCTGCGGCATACGCGTGTGCGGATTTCTTGGCAAAATATACCGCGTT 371
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 372 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGACCATTAAGAGCGTCTTTAACTCTGCA 431
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 432 GGTCTGGAAGTTCGTGAATACGCTATTATGATCGGAAATACACACTCTTGACTTCGAT 491
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 492 GCACGTGATTAAAGCCCTCAATGAAGCTCAGGCTGGCGACGTAGTGTTCATGGCTGC 551
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 552 TGGCATTAACCAACCGGTATGACCTACCTACGCTGGAACAAATGGCAAACTGGCAACTC 611
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 612 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACCAAGGTTTGGCCGCTGT 671
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaIleMetHisLysGluLeuIleVal 240
Db 672 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 731
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
Db 732 GCCAGTTCCTACTCTAAAAACTTTGGCTGTACAAACGAGCGTGTGGCGCTGTGACTCTG 791
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 792 GTTCTGCCGACAGTGAACCGTTGTATCGCGCATTCAGCAAAATGAAGCGCGGATTCGC 851
Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValVallAlaThrIleLeuSerAen 300
Db 852 GCTAACTACTCTAACCCACGACGACGCGGCTCTCTGTTGTGTCACCATCTCTGAGCAAC 911
```

```
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 912 GATCGGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCCAGCGCTATTTCAGCGT 971
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyValAenArgAspPheSerPhe 340
Db 972 AGCGGTCAAGTGTTCGTCAATACGCTGCAGAAAGGCGCAAAACCCGCACTTCAGCTTT 1031
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1032 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAAGAACAAAGTCTCGGT 1091
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1092 CTGCGCAAGAGTTCGCGTATATGCGGTTGCTTCTGTCGCTAAATGTGCGCGGATG 1151
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1152 ACACCAAGATAACATGCTCGCTGTGCGAAGCGATTGTGCGCAGTGCTG 1199

RESULT 8
ACH99107
ID ACH99107 standard; DNA; 1209 BP.
XX
AC ACH99107;
XX
XX 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polynucleotide seqid 4902.
XX
XX Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
XX Klebsiella pneumoniae.
XX
XX US6610836-B1.
XX
XX 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
XX
XX P-PSDB; ABO65556.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 4902; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX pneumoniae polypeptide of the invention
XX
XX Sequence 1209 BP; 271 A; 353 C; 343 G; 242 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 6,63e-191 Length: 1209
Score: 1931.00 Matches: 370
Percent Similarity: 96.5% Conservations: 12
Best Local Similarity: 93.4% Mismatches: 14
Query Match: 94.4% Indels: 0
DB: 11 Gaps: 0
```

US-10-673-786A-2 (1-396) x ACH99107 (1-1209)	
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20	Db 1039 ATCAGCCAGCAGAACGCGCATGTTCTCATTCAGCGCCCTGACTAAAGACGAGTGTCTGCCG 1098
Db 19 ATGTTTGAGAACATTCACCGCCGCCACCGCCGATTTAGTCTGCCGATCTGTTT 78	Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40	Db 1099 CTGGTGAAGAGTTCCGCATCTATCGGTAGCTTCCGACGCTATCAACGTGCGCGGATG 1158
Db 79 CGTCCGATACCGCCCTGAAATAATTAACTCGGAATTTGGTGTTCACAGGATGAAC 138	Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Qy 41 GlyIysThrProValLeuThrSerValLysIysAlaGluGlnTyLeuLeuGluAsnGlu 60	Db 1159 ACGCTGACATAATGGCGCCGCTGTGGAAGCCATCGTCCGCTACTG 1206
Db 139 GGTAAACCGCCGTTCTCACCAGCGTCAAAAAGCAGAGCAGTATCTCTCGAAAAATGAA 198	RESULT 9
Qy 61 ThrThrLysAsnTyLeuGlyIleAspGlyIleProGluPheGlyArgTyThrGlnGlu 80	ID ACF68649 standard; DNA; 1191 BP.
Db 199 ACGACTAAATACTATCTGGCATCATGTTTCTGTAATTTGGTCTGCCACCGAGAG 258	AC ACF68649;
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100	XX 20-NOV-2003 (first entry)
Db 259 CTGCTGTTCCGTAAGCAATCGATTATCGCTGATAACCGCCGCCACCGCGCAGACA 318	XX Photorhabdus luminescens nucleotide sequence #7116.
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120	DE XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
Db 319 CCGCGCGGTACCGGTGCGTGTCCGCGCAGACTTCTCTCGCCAAAACACCGACGTG 378	KW XX detection; food; gene expression; plant; animal; microorganism; toxin;
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140	KW XX antibiotic; biopesticide; virulence factor; disease model; plague;
Db 379 AATCGTGTGGTAAAGTAATCCGAGCTGGCGGACCATAAAGCGTATTCACTCTGCC 438	XX whooping cough; gene; ds.
Qy 141 GlyLeuGluValArgGluTyAlaTyArgAlaGluAsnHisThrLeuAspPheAsp 160	XX Photorhabdus luminescens.
Db 439 GGGCTGGAAGTGGCGAATACGATACACCGCGCTAACCGCGCTGGACTTTGAT 498	OS WO200294867-A2.
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180	PN XX 28-NOV-2002.
Db 499 GGTCTGTGGCGAGCTGCTACCGCTGCTGATTCGCTACCGCGGTGCTGTTCCACGGTGC 558	PF XX 07-FEB-2002; 2002WO-IB003040.
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyThrLeuAlaGlnLeu 200	PR XX 07-FEB-2001; 2001FR-00001659.
Db 559 TGCCACACACCGACCGGTPATCGATCCGACGCTCGATCATCGGACGCTGGCGAGCTG 618	XX (INSP) INST PASTEUR.
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220	PA XX (CNRS) CNRS CENT NAT RECH SCI.
Db 619 TCGTGGAAAAAGCTGCTACCGCTGCTGATTTCCGCTACCGAGGCTTCCGCGCGGT 678	PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
Qy 221 LeuGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240	PI Buchrieser C;
Db 679 CTGGAAGAAGATGCTGAAGGCTACCGCTTTTGTCTCTACATAAAGAGCTGCTGGTC 738	XX WPI; 2003-148459/14.
Qy 241 AlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeu 260	XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides.
Db 739 GCCAGCTCTACTCGAAAACTTTGGCTGTACACGAGCGCGTGGCGCGCTGCACTGTG 798	FS Claim 2; SEQ ID NO 7116; 1205pp; French.
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280	XX The invention relates to the isolation of genes and their encoded proteins from Photorhabdus luminescens. The isolated sequences are sources of probes and primers for detecting the genome of P. luminescens and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than P. luminescens and are able to alter response or sensitivity to toxins and antibiotics produced by P.
Db 799 GTCCCGCGGATCAGGAGCTGTAGACCGCCCTTCAGTCAAGATGAAGTCGGTATCCGC 858	CC luminescens. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically to treat microbial infection by bacteria or fungi that are sensitive to P. luminescens-encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which P. luminescens is a model (particularly plague and whooping cough). This sequence represents one of the isolated P. luminescens genes
Qy 281 AlaAsnTySerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300	CC SQ Sequence 1191 BP; 346 A; 279 C; 267 G; 299 T; 0 U; 0 Other;
Db 859 GCCAACTACTCGAACCCCGCTCGCATGGCGCTCCGTTGGTGGTCCACCATTTCTGAGCAAC 918	Alignment Scores:
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320	
Db 919 GATCGCTACGGGCAATCTGGGAGCAGGAACGTACCGATATCGCGCAGCGATCCAGCGT 978	
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe 340	
Db 979 ATGGCTGTGCTGTTCGTCAATACCTTCAGGAGAAAGCGCGAGCGGATTCAGCTTT 1038	
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360	

```
Pred. No.: 5.45e-165 Length: 1191
Score: 1682.00 Matches: 316
Percent Similarity: 89.4% Conservative: 38
Best Local Similarity: 79.8% Mismatches: 42
Query Match: 82.2% Indels: 0
DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x ACF68649 (1-1191)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAAATCACCAGCGCGCTGCCGACCTATTCTTGGCTTAGCCGATGTTTC 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 CGTTTCGATCCTCTACAAATCAAAATCAACCTTGGTATCGGTGTCTATAAAGACGAACA 120
Qy 41 GlyIleThrProValLeuThrSerValLysIleAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 121 GGAATAACCCCGGTTCTGACGAGTGTAAAGAGCTGAACAATATTTACTGGAAACGAA 180
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACAACAAGATATATCTGCCGATTAGCGGCTTAGCCGAATTTGGCCGCGTAATCAAGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 TTTACTGTTTGGCAAGATCACCAGCTGTGTACAGATAAAGCGCGCCGACAGCACAAAGC 300
Qy 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CCAGGCGGATCCCGGTGCTTTACGATTGCTGCGGATTTTATGTCGCAACAGACTAATGCT 360
Qy 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAACAGATTGGATCAGCAACCAACCTGGCCAAACCACTAAATTCGAA 420
Qy 141 GlyLeuGluValArgGlyTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTCTGTGAATATAAATACTATGACGCTGAAACACGCGCTGAATTCGAA 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GACATGTGCGCAACGCTGTCGGAAGCTCAGCTGTGTGTCTGTTCCACGCGCTGC 540
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 541 TGCCACAAATCCGACAGCGATCGATCCAAACCCGCGCACAAATGGGCTAACTGGCAGAAATG 600
Qy 201 SerValGlyLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCTCGGAGAAAGGCTGGTTGCTATTTTGGTATTCGCTTACCAAGGATTTGCCAAGGC 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 CTAAACGAAGATGACAGAGGCGCTACGTATTTTTCGGAATAATCATATGAAGTATGTT 720
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValAlaCysThrLeu 260
Db 721 GCCAGCTCTTACTCCAAANAACTTTGGCTGTACAAATGAAGTGTGCGTGTGCTGACTATT 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGCTAGTGACAGTATACAGCAAGAAAGCGTTACGCAAGCTAAAGCGATATCCGT 840
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 GCTAACTATTCCAAACCCCGCTCATGGTGCATCTATTGTCTACATACATTTTGTCAAT 900
Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GAAGACCTGAAGAGCGCTTGGGAACAGAGAACTGACACCATCGCGCAACGTATCCAGCGT 960
```

```
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 961 ATCGGTCAACTGTTTGTCAATACCTTGCAGAAAGGGGCAAAACAGGATTTTCAGCTTT 1020
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATTATTAGCCAAATGGTATGTTCTCATTCAGTGGCTGACAAAAGAACAGTAGAACGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 CTGCGTGATGAGTTGGTATATACGCTGTACGTTCCGTCGTATTAACTGCTGGCTTG 1140
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACGTTAGAAAACATGGCTCCACTATGTGAAGCCATTTGTCGAGTACTC 1188

RESULT 10
ACF67367_14
Continuation (15 of 57) of ACF67367 from base 1400001 (Photorhabdus luminescens nucleoti
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
WP Fragment Name Begin End
WP ACF67367_00 1 110000
WP ACF67367_01 100001 210000
WP ACF67367_02 200001 310000
WP ACF67367_03 300001 410000
WP ACF67367_04 400001 510000
WP ACF67367_05 500001 610000
WP ACF67367_06 600001 710000
WP ACF67367_07 700001 810000
WP ACF67367_08 800001 910000
WP ACF67367_09 900001 1010000
WP ACF67367_10 1000001 1110000
WP ACF67367_11 1100001 1210000
WP ACF67367_12 1200001 1310000
WP ACF67367_13 1300001 1410000
WP ACF67367_14 1400001 1510000
WP ACF67367_15 1500001 1610000
WP ACF67367_16 1600001 1710000
WP ACF67367_17 1700001 1810000
WP ACF67367_18 1800001 1910000
WP ACF67367_19 1900001 2010000
WP ACF67367_20 2000001 2110000
WP ACF67367_21 2100001 2210000
WP ACF67367_22 2200001 2310000
WP ACF67367_23 2300001 2410000
WP ACF67367_24 2400001 2510000
WP ACF67367_25 2500001 2610000
WP ACF67367_26 2600001 2710000
WP ACF67367_27 2700001 2810000
WP ACF67367_28 2800001 2910000
WP ACF67367_29 2900001 3010000
WP ACF67367_30 3000001 3110000
WP ACF67367_31 3100001 3210000
WP ACF67367_32 3200001 3310000
WP ACF67367_33 3300001 3410000
WP ACF67367_34 3400001 3510000
WP ACF67367_35 3500001 3610000
WP ACF67367_36 3600001 3710000
WP ACF67367_37 3700001 3810000
WP ACF67367_38 3800001 3910000
WP ACF67367_39 3900001 4010000
WP ACF67367_40 4000001 4110000
WP ACF67367_41 4100001 4210000
WP ACF67367_42 4200001 4310000
WP ACF67367_43 4300001 4410000
WP ACF67367_44 4400001 4510000
WP ACF67367_45 4500001 4610000
WP ACF67367_46 4600001 4710000
WP ACF67367_47 4700001 4810000
WP ACF67367_48 4800001 4910000
WP ACF67367_49 4900001 5010000
WP ACF67367_50 5000001 5110000
WP ACF67367_51 5100001 5210000
```


WP AC67367_52 5200001 5310000
 WP AC67367_53 5300001 5410000
 WP AC67367_54 5400001 5510000
 WP AC67367_55 5500001 5610000
 WP AC67367_56 5600001 5648894

Alignment Scores:
 Pred. No.: 2,97e-162 Length: 110000
 Score: 1682.00 Matches: 316
 Percent Similarity: 89.4% Conservatives: 38
 Best Local Similarity: 79.8% Mismatches: 42
 Query Match: 82.2% Indels: 0
 DB: Gaps: 0

US-10-673-786A-2 (1-396) x ACF67367_14 (1-110000)

Qy	1	MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGluValAlaAspLeuPhe	20
Db	23013	ATGTTTGAGAAATACACCGAGCGCTTCCGACCCCTATTCTTGGCTTAGCCGATAGTTTC	23072
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr	40
Db	23073	CGTTCTGATCCTCGTACAAATAAATCAACCTTGGTATCGGTGCTATATAAGACGAAACA	23132
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu	60
Db	23133	GGAAAAACCCGGTCTCTGACCACTGTATAAAAGCTGAACAATATTACTTGGAAAAACGAA	23192
Qy	61	ThrThrLysAsnTyrLeuGlyLysIleAspGlyLysProGluPheGlyValGlnGlu	80
Db	23193	ACAACAAGAAATATTCTGCGGATAGCGGCTTAGCCGAATTTGGCGCGCTAACTCAAGAA	23252
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	23253	TTACTGTTGGCAAGATACACCACTGTGTACAGATAAACGCCCGCCGACAGACAAAGC	23312
Qy	101	ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	23313	CCAGGCGGTACCGGTGCTTTACGTTATTGCTGCGGATTTTCATTTGCCAAACAGACTAATGT	23372
Qy	121	LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
Db	23373	AAACGAGTTTGGATCAGCAACCCCACTGCCAAACCAATAAAACGCTTTTTCGCGCGCT	23432
Qy	141	GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	23433	GGTCTGGAAGTCTGTAATATAATAACTATGACGCTGAAAAACACGCGCTGATTTGCA	23492
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	23493	GACATGTGCGCAAGCCTGTCCGAAGCTCAGGCTGGTGATGTTGTTCTGTTCCACGGCTC	23552
Qy	181	CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrThrLeuAlaGlnLeu	200
Db	23553	TGCCCAATCCGACGACATCGATCCAAACCCCGSCCAATGGCTTAAACTGGCAGAAATG	23612
Qy	201	SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
Db	23613	TCGCGGAGAAAGCGTGTGGCTTATTTTGGATTTCCGATGATTTCCAGGATTTCCAGGCG	23672
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuVal	240
Db	23673	CTAAACGAGATGCAGAGGCGCTACGTTATTTTGGCAAAATCATATGACTGATGTTT	23732
Qy	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
Db	23733	GCCAGCTCTTACTCCAAAACTTTGGCCTGTACAATCAACGATGTCGGTCCGCTGACTATT	23792
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
Db	23793	GTTGCTAGTGACAGTGATGATCAGCAGAAAAACCGCTTCCAGCAGCTAAACGATTAATCCG	23852
Qy	281	AlaAsnTyrSerAsnProProAlaHisGlyValSerValValAlaThrIleLeuSerAsn	300

CC recombinant production of the proteins, particularly toxins and
 CC antibiotics useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 249878 BP; 75946 A; 56652 C; 49249 G; 68029 T; 0 U; 2 Other;
 Alignment Scores:
 Pred. No.: 9.3e-162 Length: 249878
 Score: 1682.00 Matches: 316
 Percent Similarity: 89.4% Conservative: 38
 Best Local Similarity: 79.8% Mismatches: 42
 Query Match: 82.2% Indels: 0
 DB: 10 Gaps: 0
 US-10-673-786A-2 (1-396) x ACF65381 (1-249878)
 QY 1 MetPheGluAenIleThrAlaAProAlaAAspProIleLeuGlyLeuAlaAAspLeuPhe 20
 DB 1334 ATGTTTGAGAAATACCGCAGCGCTGCCGACCTATTCTTGGCTTAGCCGATAGTTTC 1393
 QY 21 ArgAlaAAspGluAArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAAspGluThr 40
 DB 1394 CGTTCGTGATCTCGTACAAATATAATCAACCTTGGTATCGGTGCTATATAGAGCAAAACA 1453
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu 60
 DB 1454 GGAATAACCCCGGTTCTGACCAGGTGTTAAAAAGCTGAACAATATTTACTGGAAAAACGAA 1513
 QY 61 ThrThrLysAenTyrLeuGlyIleAAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 1514 ACAACAAAGAATATTATGCCGATTAGCGGCTTAGCCGAATTTGGCGGCTTAACCTCAAGAA 1573
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAArgThrAlaGlnThr 100
 DB 1574 TTACTGTTTGCAAGATACCCAGTGTGTACAGATAAACCGCGCCGACAGCAACAAGC 1633
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAAspPheLeuAlaLysAenThrSerVal 120
 DB 1634 CCAGCGGTACCGGTGCTTTACGTATTGCTGCGGATTTCAATGGCAACAGACTAATGCT 1693
 QY 121 LysArgValTropValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 DB 1694 AACCGAGTTTGGATCAGCAACCAACCTGGCCAAACCATATAAAACGTTTTTTCGCGCGCT 1753
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAAspPheAAsp 160
 DB 1754 GGTCTGGAGTCTGTGAATATAAATATATATGACCTGAAACACAGCGCTGAATTTGAA 1813
 QY 161 AlaLeuIleAenSerLeuAAsnGluAlaGlnAlaGlyAAspValValLeuPheHisGlyCys 180
 DB 1814 GACATGCTGGCAAGCGCTGCCAAGCTCAGCTGGTGTATGTTGTTCTTCCACGGGTGC 1873
 QY 181 CysHisAenProThrGlyIleAAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 DB 1874 TGCCACAATCCGACAGGCGATCGATCCACCCCGGCAACATGCGGTAACTGGCAGAAATG 1933
 QY 201 SerValGluLysGlyTropLeuProLeuPheAAspPheAlaTyrGlnGlyPheAlaAArgGly 220
 DB 1934 TCTCGGAGAAAGCGTGGTTCCTATTGTTGATTTGCTTACCAAGGATTTGCCAAGGCG 1993
 QY 221 LeuGluGluAAspAlaGluGlyLeuArgAlaPheAlaAalaMetHisLysGluLeuIleVal 240
 DB 1994 CTAACCAAGATGCAGAGGCGCTACGTATTTTTCGAAAAATCATATGAATGAATCATTTGTT 2053
 QY 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
 DB 2054 GCCAGCTCTTACTCCAAAAAATTTTGGCGCTGTACAAATGAACGTGTCGGTGGCTGTACTATT 2113

QY 261 ValAlaAlaAAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 DB 2114 GTTGTAGTACAGTGTATACAGCAGAAAAAGCGTTACGCCAAGCTAAAGCGATTATCCGT 2173
 QY 281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 DB 2174 GCTAACTATTCCAAACCCACCGCTCATGGTGCATCTATTGTCTACTACATTTTGTCAAT 2233
 QY 301 AspAlaLeuArgAlaIleTropGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 2234 GAAGACCTGAAGACGAGCTTGGGAACAGGAACCTGACCACCATGCGCAACGATATCCAGCGT 2293
 QY 321 MetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAAspPheSerPhe 340
 DB 2294 ATGCGTCAACTGTTTGTCAATACCTTGCAGGAAAAAGGGCAAAACAGGATTTTCAGCTTT 2353
 QY 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 2354 ATTATTAGCCAAAATGGTATGTTCTCATTCAGTGGCTGCACAAAGAACAGTAGAACGT 2413
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
 DB 2414 CTGCGTGATGAGTTTGGTATATACGCTGTCAGTTCGGTCTGTTAAACGTCGCTGGCTTG 2473
 QY 381 ThrProAAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 2474 ACGTAGAAAACATGGCTGCACCTATGTGAAGCCATTGTTGCAGTACTC 2521
 RESULT 12
 ADT46374
 ID ADT46374 standard; cDNA; 1188 BP.
 XX
 AC ADT46374;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #21125.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 4812; 122pp; English.
 XX

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 1188 BP; 340 A; 268 C; 269 G; 311 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,478-162 Length: 1188
 Score: 1654.00 Matches: 310
 Percent Similarity: 86.9% Conservative: 34
 Best Local Similarity: 78.3% Mismatches: 52
 Query Match: 80.9% Indels: 0
 DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADT46374 (1-1188)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 1 ATGTTGAAATAATACAGCAGCAGCTGCCGACCTATCTCTGTTAGCGGATAGCTTT 60
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 61 AAAGCTGATCTCGTGAAATAAATAATCAATCTGGGTATCGGTGTTTACAAAGACGAAACT 120
 Qy 41 GlyIysThrProValLeuThrSerValIysIysAlaGluInTyrLeuLeuGluAsnGlu 60
 Db 121 GGTAACACCCCTGCTGACACCCGTTAAATAAGCAGAAAAATTCCTGCTGGAACCGAA 180
 Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 181 ACCACCAATAATATCTGGCAATTAAGTGGTACCTGAATTTGGCGGCTACTCAGGAA 240
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 241 CTACTCTTCGGCAATACACGAGCTATCATTTACTGACAAACGCGCCGTACTGTACAAAGC 300
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 301 CCAGCGGAAACCGGTGCACTGCGCACTCCGCGAGATTTTATTTGCTTAAACAGACCAATGCA 360
 Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 361 AAACGTGTCTGATCAGTAACCAACATGGCCAAACCAATTAAGGCGTTTCTTAGCGCT 420
 Qy 141 GlyLeuGluValArgGlyTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 421 GGGTTAGAAATCCCGAGTACCAATATTACAAATGACGAAAGCATGCTCTGATTTTCAT 480
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GGCATGTTGGTCTGCTGTAAGCTCAAGCAGGTGATGTTGTTCTGCTTCAACGCTTGC 540

Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 541 TGCCATAACCCCAACCGGTATCGATCCAAACGCTGAACAGTGGCAGAAACTTGCAGATTTA 600
 Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 601 TCTCGGCAAAATGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 661 TTAGATGAAGATGACGAAGGCTGCGTATTTTACTAAACCAATAAGTATGATCGTT 720
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGCTTCTACTCCAAAATTTTGGCTGTACAAATGACGCTGTAGGTGCTGCCACCATC 780
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGCAACTGACAGCGATACCGCAGAAAGCAATTCAGTCAGCGCAAAATCTATTGTTCT 840
 Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaIleThrIleLeuSerAsn 300
 Db 841 ACTAACTACTCTAACCCACCGCACCGCGCTTCTGTTGTACTACAAATTTCTGTCAT 900
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 901 GACGAATTTAAAGCAGAAATGATTCAGAACTGCGCACTATCGGTGAACGCAATTCGTC 960
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspSerPhe 340
 Db 961 ATGCGTCAGCTGTTTGTGAACACCTTGCAAGAAAGAGGTGCAAAACAGACTTTAGCTTT 1020
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATTATTACCAAAATGATGTTTTCATTCAGTGTCTGACTAAAGAACAGTTGATCGT 1080
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGTGAAGAAATTCGGTATTTATGCGGTGCTGCGGTGCTGCGGTGCTGCGGTGCTG 1140
 Qy 381 ThrProAsnAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACCTGGAGAACATGGTTCTCTGTTGTAAGCGGATTTGTCAGTACTC 1188

RESULT 13
 ID ADT05500 standard; DNA; 7977 BP.
 XX
 AC ADT05500;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 536.
 XX
 KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
 XX
 OS Haemophilus influenzae.
 XX
 PN WO2004078949-A2.
 XX
 PD 16-SEP-2004.
 XX
 PF 05-MAR-2004; 2004WO-US007001.
 XX
 PR 06-MAR-2003; 2003US-0453134P.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL INC.
 XX
 PI Bakaletz LO, Munson RS, Dyer DW;
 XX
 DR WPI; 2004-662422/64.
 XX

PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
 PT useful for treating or preventing NTHi bacterial infections of the middle
 PT ear and/or nasopharynx.
 XX
 PS
 PS Example 1; SEQ ID NO 536; 88pp; English.

CC The invention comprises nucleotide sequences (genes) from the genome of a
 CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
 CC sequences of the invention are useful for treating or preventing NTHi
 CC bacterial infections of the middle ear and/or nasopharynx. The present
 CC nucleic acid represents an NTHi contig sequence of the invention.
 XX
 XX Sequence 7977 BP; 2664 A; 1592 C; 1464 G; 2255 T; 0 U; 2 Other;

Alignment Scores:
 Pred No.: 1,51e-128 Length: 7977
 Score: 1343.00 Matches: 251
 Percent Similarity: 78.3% Conservative: 59
 Best Local Similarity: 63.4% Mismatches: 86
 Query Match: 65.7% Indels: 0
 DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADT05500 (1-7977)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 1594 ATGTTGAACATATTAAAGCGGACCGGATCCATCTTAGGCGTAGGCGAGCATTC 1653
 Qy 21 ArgAlaAspGluAargProGlyLysIleAsnLeuGlyLeuGlyValTyrLysAspGluThr 40
 Db 1654 AAATCCGAACCTCGGAAAAATAAAATTAACCTGGGTATGGCGTTTATAAAGATGCCCAA 1713
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 1714 GGCACACCCGATATGCGTGCAGTGAAGACGACGAGAAACCGCTTATTGTAAAGGAA 1773
 Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyA:GcysThrGlnGlu 80
 Db 1774 AAAACGAAAAATTATCTGACTATCGATGATGCTGATGATGATGATGATGATGATGATG 1833
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 1834 TTACTTTTCGGTAAAGATTCTGAAGTATCATCAATCTAATCGTGCAGAACACAGTACAAAGT 1893
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 1894 TTAGCGGACAGCGGATACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1953
 Qy 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
 Db 1954 CAAATGTTTGGATCAGCAGCCCACTTGGCCAAACACCAATGCGATTTTCAATGCGGTC 2013
 Qy 141 GlyLeuGluAlaArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 2014 GGTATGCAATTCGTGATATCGTATTATGATGCTGAACGCAAAAGCCCTTTGATGGGAT 2073
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2074 CATTACTTGAAGATTTAAGCAAGCAAGCGAAGCGATGCGTGTTCACCGTTGT 2133
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 2134 TGCCATAATCCGACTGGTATTGACCTTACTCCAGAACAAATGGCAAGATTAAGCGCACTT 2193
 Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 2194 TCGCGGAAAAACGGCTGGTGGCACTCTTGACTTTGCTTATCAAGTTTAGGCAATGGA 2253
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 Db 2254 TTAGATGAGATGCTTATGGCTTGGCTTTGGCGCAACCAACCAAGATTTGTAGTA 2313
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260

Db 2314 GCAGTTCATTCTCGAAAAAAGCTTTGGTTATATATGAACGTTGGTGGCTTACTCTT 2373
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 2374 GTGGCAGAAAATGCAGAAATTCATCAACCGCATTAACCCCAAGTGAATCAATATTTCGT 2433
 Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 2434 ACACCTTACTCTTAACCTCGATCTACGGCGGCGCAACTGTAGCAACAGTATTAAATGAC 2493
 Qy 301 AspAlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 2494 ACTCAACTTCGCCAAGAGTGGGAAATGAATTAACCTGAATGCGGCAACGATCAAAAA 2553
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPhe 340
 Db 2554 ATCGGTCAATTATTCGTTTCAGTTATTAAAGAAATACGGTCGGGCAACAGATTTTCAGCTTT 2613
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 2614 ATCATTGAACAAAAACGATATGTTCTCTTTCAGTGGATTAACTGGGGAACAGTGATCGC 2673
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 2674 TTAAGAAGAAATTTGCCATTATGCTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTC 2733
 Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 2734 ACTGAAGATTAATTCGCTATTATGTGAAGCATTGTGAAGTGTCTT 2781

RESULT 14

ADT05649/c
 ID ADT05649 standard; DNA; 349980 BP.

XX AC ADT05649;

XX DT 02-DEC-2004 (first entry)

XX DE Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 685.

XX DE middle ear bacterial infection; nasopharynx bacterial infection; ds.

XX OS Haemophilus influenzae.

XX PN WO2004078949-A2.

XX PD 16-SEP-2004.

XX PF 05-MAR-2004; 2004WO-US007001.

XX PR 06-MAR-2003; 2003US-0453134P.

XX PA (CHIL-) CHILDRENS HOSPITAL INC.

XX PI Bakaletz LO, Munson RS, Dyer DW;

XX DR WPI; 2004-662422/64.

XX PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
 XX useful for treating or preventing NTHi bacterial infections of the middle
 XX ear and/or nasopharynx.

XX PS Claim 1; SEQ ID NO 685; 88pp; English.

XX CC The invention comprises nucleotide sequences (genes) from the genome of a
 XX nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
 XX sequences of the invention are useful for treating or preventing NTHi
 XX bacterial infections of the middle ear and/or nasopharynx. The present
 XX nucleic acid represents an NTHi DNA sequence of the invention.

XX SQ Sequence 349980 BP; 105127 A; 63538 C; 70035 G; 111276 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 2,92e-126 Length: 349980
 Score: 1343.00 Matches: 251
 Percent Similarity: 78.3% Conservative: 59
 Best Local Similarity: 63.4% Mismatches: 86
 Query Match: 65.7% Indels: 0
 DB: 13 Gaps: 0

US-10-673-786A-2 (1-396) x ADT05649 (1-349980)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGluLeuAlaAspLeuPhe 20
 DB 217833 ATGTTGAACATATTAAAGCGGCACAGCCGATCCAAATCTTAGGCTTAGGCGAAGCATTC 217774
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
 DB 217773 AAATCCGAAACTCGGAAATTAATAATTAATCTGGGTATGGCTTTAAGATCGGCAA 217714
 QY 41 GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyLLeuGluAsnGlu 60
 DB 217713 GGCACACCCCGATTATGCTGCAGTGAAGAAGCAGAAAAACGGCTTATTGTGATAAGGAA 217654
 QY 61 ThrThrLysAsnTyLLeuGlyLysIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 217653 AAAACGAAAAATTTATCTGACTATCGATGCTGTTATTAACGAAACAAACAAAGCG 217594
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysAlaArgThrAlaGlnThr 100
 DB 217593 TTACTTTTCGGTAAGATTCTGAAGTCATCAATCTAATCTGTCGCAAGAACAGTACAAAGT 217534
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 217533 TTAGCGGGAACAGCGCATTCAGCATTCGCGCAGAAATTTATTAACGTCAAACAAAGCA 217474
 QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 DB 217473 CAANAATGTTGGATCAGCACCCCACTTGGCCAAAACCAATGCCATTTTCAATCCGCTC 217414
 QY 141 GlyLeuGluValArgGluTrpAlaTyLysAspAlaGluAsnHisThrLeuAspPheAsp 160
 DB 217413 GGTATGACAAATTCGTGAATATCGTTATTATGATCTGAACGCAAGCCCTTGATGGGAT 217354
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 DB 217353 CATTTACTTGAGATTAAAGCCAGCAAGCGAAGCGATGTGCTTTTACACGGTTGT 217294
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 DB 217293 TGCCATAATCCGACTGGTATTGACCCCTACTCCAGAACCAATGGCAAGAATTTAGCCGACTT 217234
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyLLeuGlnGlyPheAlaArgGly 220
 DB 217233 TCGCGGAAAACCGGTGGTGGCATCTTTGACTTTGTCTTATCAAGGTTTAGCCCAATGGA 217174
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 DB 217173 TTAGATGAAGATGCTTATGCTTGGCTTGGCTTTGGCGCAACCAACAGAATTTAGTAGTA 217114
 QY 241 AlaSerSerTyLLeuAsnPheGlyLeuTyLLeuAsnGluArgValAlaGlyCysThrLeu 260
 DB 217113 GCGAGTTCATCTCGAAAAAATTTGGTTTATATGAACGTTGTTGGTGGCTTTACTCTT 217054
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaLeuArg 280
 DB 217053 GTGCGAGAAAATTCGAGAAATTCGATCAACCGCATTAACCCCAAGTGAATTTATTCGT 216994
 QY 281 AlaAsnTyLLeuSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 DB 216993 ACATCTTACTTAACCTGTCATCTCAGCGCGCGCAACTGTAGCAACAGTATTAAATGAC 216934
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
 DB 216933 ACTCAACTTCGCAAGAGTGGGAAAAATGAATTAATCTGAATTCGCGCAACGTTATCAAAAAA 216874

QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 DB 216873 ATGGCTCATTTATTTCAGTTATTAAAGAATAATACGCTGGCGAAACAAGATTTCAGCTTT 216814
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 DB 216813 ATCATTTGAACAAAACGGTATGTTCTTTTCAGTGGATTAACTGGGGAACAAGTGGATCGC 216754
 QY 361 LeuArgGluGluPheGlyValTyLLeuAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 DB 216753 TTAAAGAGAAATTTGCCATTATGCTGTTCTGTTCTGCTGATCAATGCTGCTGATTT 216694
 QY 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 DB 216693 ACTGAAGATAATATTTCGCTATTATTTATGTAAGCATTTGTGAAAGTCTTT 216646

RESULT 15

AAT42063_16

Continuation (17 of 19) of AAT42063 from base 1600001 (Haemophilus influenzae complete
 WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000
WP	AAT42063_10	1000001	1110000
WP	AAT42063_11	1100001	1210000
WP	AAT42063_12	1200001	1310000
WP	AAT42063_13	1300001	1410000
WP	AAT42063_14	1400001	1510000
WP	AAT42063_15	1500001	1610000
WP	AAT42063_16	1600001	1710000
WP	AAT42063_17	1700001	1810000
WP	AAT42063_18	1800001	1830121

Alignment Scores:

Pred. No.: 4.37e-125 Length: 110000
 Score: 1325.00 Matches: 248
 Percent Similarity: 77.8% Conservative: 60
 Best Local Similarity: 62.6% Mismatches: 88
 Query Match: 64.8% Indels: 0
 DB: 2 Gaps: 0

US-10-673-786A-2 (1-396) x AAT42063_16 (1-110000)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 DB 84147 ATGTTGAACATATTAAAGCGGCACAGCCGATCCAAATCTTAGGCTTAGGCGAAGCATTT 84206
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
 DB 84207 AAATCCGAAACTCGGAAATTAATAATTAATCTGGGTATGGCTTTAAGATCGGCAA 84266
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLLeuGluAsnGlu 60
 DB 84267 GGCACACCCCAATTTATGACCGCGTAAAGAAGCGGAAACGATTTATTGTGATAAGGAA 84326
 QY 61 ThrThrLysAsnTyLLeuGlyLysIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 84327 AAAACCAAGAATTTATCTGACTATCGATGCTGTTATTCGCGGATTAACCAACAAACAAAGCA 84386
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysAlaArgThrAlaGlnThr 100
 DB 84387 CTCCTTTTCGGTAAAGATTCTGAAGTCATCCATCTAATTCGAGCAAGAACAGTACAAAGT 84446
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120

[illegible]

Search completed: March 15, 2006, 03:34:59
Job time : 1478 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:18:01 ; Search time 5196 Seconds
(without alignments)
4332.180 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045
Sequence: 1 MFENITAAPDPIILGLDLF.....VAGMTDPDNAPLCEAIVAVL 396

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB/spool/US10673786/runat_14032006_135609_22591/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -HOST=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03p
-USER=US10673786_QCGN_1_1_7415_@runat_14032006_135609_22591 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

RESULT 1

108485
LOCUS
DEFINITION Sequence 12 from Patent WO 8700202.
ACCESSION I08485
VERSION I08485.1 GI:588805
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1293
/organism="unknown"
/mol_type="unassigned DNA"

Result No. Score Match Length DB ID Description

1 2045 100.0 1293 6 I08485
2 2045 100.0 1331 1 ECASPCG
3 2045 100.0 1331 6 AR095441

SUMMARIES

Alignment Scores:
Pred. No.: 7.8e-149 Length: 1293
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

2045 100.0 1331 6 E16758
2045 100.0 1415 1 ECASPC
2045 100.0 1468 1 I08484
2045 100.0 3659 6 I08487
2045 100.0 11470 1 D90731
2045 100.0 14759 6 AX370270
2045 100.0 20604 1 D90730
2045 100.0 110000 1 U00096_09
2037 99.6 110000 1 AE005174_11
2037 99.6 110000 1 AE005174_12
2037 99.6 110000 1 BA000007_11
2033 99.4 301276 1 AE016758
2026 99.1 110000 1 AE005674_09
2026 99.1 292504 1 AE016981
1982 96.9 62027 1 AE008743
1979 96.8 110000 1 AE017220_10
1978 96.7 110000 1 CP000026_18
21 1978 96.7 294050 1 AL627268
22 1978 96.7 301983 1 AE016840
23 1931 94.4 1209 6 AR388173
24 1780 87.0 110000 1 BX950851_28
25 1775 86.8 14855 1 AE013879
26 1775 86.8 110000 1 BX936398_17
27 1775 86.8 210050 1 AJ414148
28 1775 86.8 291847 1 AE017131
29 1682 82.2 249878 6 AX770903
30 1682 82.2 349071 1 BX571864
31 1349 66.0 9431 1 AY795600
32 1343 65.7 7977 6 CQ872923
33 1343 65.7 110000 1 CP000057_13
34 1343 65.7 349980 6 CQ873072
35 1336 65.3 110000 1 CP000020_16
36 1333 65.2 348672 1 CR378670
37 1329 65.0 10952 1 U32835
38 1325 64.8 10766 1 AE015676
39 1325 64.8 110000 1 BA000031_19
40 1325 64.8 110000 6 BD426631_16
41 1325 64.8 110000 6 AR274513_16
42 1325 64.8 110000 6 AF632719_16
43 1324 64.7 7831 1 AF183931
44 1318 64.4 300029 1 AE016804
45 1317 64.4 110000 1 AE004969_14

ALIGNMENTS

linear PAT 02-DEC-1994

1 (bases 1 to 1293)
Edwards, M.R., Taylor, P.P., Hunter, M.G. and Fotheringham, I.G.
COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
Patent: WO 8700202-A 12 15-JAN-1987;
Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

US-10-673-786A-2 (1-396) x 108485 (1-1293)	
Qy	1 MetPheGluAenIleThrAlaIaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db	12 ATGTTTGAGAACATTACCGCGCTCTTCCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 71
Qy	21 ArgAlaAspGluAArgProGlyIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db	72 CGTGCCGATGAACCTCCCGGCAAAATTAACCTCGGGATGGTGCTATAAGATGAGACG 131
Qy	41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu 60
Db	132 GGCAAAACCCCGGTACTACACAGCGTGAAAAGGCTGACACGATATCTGCTCGAAATGAA 191
Qy	61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyAArgCysThrGlnGlu 80
Db	192 ACCACCAAAATTAACCTCGGCATTCAGCGCATCCCTGAATTTGGTCTGCACTCAGAA 251
Qy	81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaAArgThrAlaGlnThr 100
Db	252 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCACGGCCACAGACT 311
Qy	101 ProGlyGlyThrGlyAlaLeuAArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db	312 CCGGGGGCACTGGCGGCATACGCTGGCTGCCATTTCTTGGCAAAAATACACGCGTT 371
Qy	121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAAsnSerAla 140
Db	372 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGAACCAATAAGACGCTCTTTAACTTGCA 431
Qy	141 GlyLeuGluValAArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db	432 GGTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAAATCACACTCTTGACTTCGAT 491
Qy	161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db	492 GCACGTGATTACACGCTGAATGAAGCTCAGGCTGGCAGCGTAGTGTCTTCATGGCTGC 551
Qy	181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db	552 TGCATAACCCCAACCGGTATCGACCCCTACGCTGCGAACAATGGCAAAACACATGGCAACATC 611
Qy	201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaAArgGly 220
Db	612 TCCGTTGAGAAAGGTGTTTACCCTGTGTTGACTTCGCTTACCAAGGGTTTTGCCCGTGGT 671
Qy	221 LeuGluGluAAspAlaGluGlyLeuAArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db	672 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCATTAAGAGACTGATTGTT 731
Qy	241 AlaSerSerTyrSerLysAAsnPheGlyLeuTyrAenGluAArgValGlyAlaCysThrLeu 260
Db	732 GCCAGTTCTTACTCTAAAACTTTGGCCTGTACACGAGCGTGTGGCGCTTTGTACTCTG 791
Qy	261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleAArg 280
Db	792 GTTGCTGCGCAGTGAACCGTTGATCGCGCATTCAGCAAAATGAAGCGCGCATTCGC 851
Qy	281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db	852 GCTAACTACTTAACCCACAGCACACGCGCTTCTGTGTGGCCACCATCCTGAGCAAC 911
Qy	301 AspAlaLeuAArgAlaIleTrpGluGlnGluLeuThrAAspMetArgGlnArgIleGlnArg 320
Db	912 GATGGTTACGTGGCATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGGT 971
Qy	321 MetArgGlnLeuPheValAAsnThrLeuGlnGluLysGlyAlaAAsnArgAspPheSerPhe 340
Db	972 ATGCGTCAGTTGTTTCGTCAATACGTCGAGGAAAAAGCGCAAAACCCGCACTTCAGCTTT 1031
Qy	341 IleIleLysGlnAAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuAArg 360
Db	1032 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGCACAAAAGAACAAAGTGTCTCGT 1091
Qy	361 LeuAArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db	1092 CTCGCGGAGAGTTTGGCGTATATGCGGTTCCTTCGTCGCTAAATGTTGGCGCGGATG 1151
Qy	381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db	1152 ACACCAAGATAACATGCTCGCTGTGGAAGCGATTTGGCACTGTGCTG 1199
RESULT 2	
ECASPCG 1331 bp DNA linear BCT 12-SEP-1993	
LOCUS E. coli aspC gene for aspartate aminotransferase.	
DEFINITION X05904.1 GI:41012	
ACCESSION aspartate aminotransferase; aspC gene.	
VERSION Escherichia coli	
KEYWORDS Escherichia coli	
SOURCE Escherichia coli	
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.	
REFERENCE 1 (bases 1 to 1331)	
AUTHORS Kuramitsu, S., Okuno, S., Ogawa, T., Ogawa, H. and Kagamiyama, H.	
TITLE Aspartate aminotransferase of Escherichia coli: nucleotide sequence of the aspC gene	
JOURNAL J. Biochem. 97 (4), 1259-1262 (1985)	
PUBMED 3897210	
REFERENCE 2 (bases 1 to 1331)	
AUTHORS Kagamiyama, H.	
TITLE Direct Submission	
JOURNAL Submitted (07-MAR-1988)	
COMMENT Data kindly reviewed (07-MAR-1988) by Kagamiyama H.	
FEATURES	
source	1..1331
	/organism="Escherichia coli"
	/mol_type="genomic DNA"
	/db_xref="taxon:562"
	10..1200
	/note="unnamed protein product; aspartate aminotransferase (AA 1 - 396) "
	/codon_start=1
	/transl_table=11
	/protein_id="CAA29333.1"
	/db_xref="GI:41013"
	/db_xref="GOA:P00509"
	/translation="MFENITAAPADPILGLADLFRADRPCKINLIGIVKDETKTP VLTSVKAEQYLLENETTKNYLGIDGIPFGRCTQELLFGKGSALINDKRAARTQTPG
	ALINSNEAQAGDVLPFHGCCNPTGIDPTLEOMOTLAQLSVEKGMPLPDPAYQGFAR
	RGLEBDAEGLRAPAAHKEILIVASSYKNGFLYNERVGACTLVAAADSETVDPAFSOMK
	AALRANYSNPPAHGASVYATILNSDALRALIWEQELTDMRQRIQMRQLFVNTLQEKGA
	NRDPSFIIKQNGMFSFGLTKQVLRLEBEFVYAVASGRVNVAGMTPDNNAPLCEAI
	VAVL"
old_sequence	90
	/note="c was a in [1]"
	/citations=[1]
old_sequence	502..503
	/note="ag was tc in [1]"
	/citations=[1]
old_sequence	612
	/note="c was g in [1]"
	/citations=[1]
old_sequence	657
	/note="t was g in [1]"
	/citations=[1]
old_sequence	693
	/note="g was c in [1]"
	/citations=[1]
old_sequence	696
	/note="c was t in [1]"
	/citations=[1]
old_sequence	699

/notes="t was g in [1]"
/citation={1}

ORIGIN

Alignment Scores:
Pred. No.: 8.08e-149 Length: 1331
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x ECASPCG (1-1331)

```
QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 10 ATGTTTGAGAACATTACCGCGCTCTCGGACCGCATTCGGCCCTGGCCGATCTGTTT 69
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
DB 70 CGTGGCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 129
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
DB 130 GGCACAAACCCCGCTACTGACCGCTGAAAGAGGCTGAACAGTATCTGCTCGGCGATG 189
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 190 ACCACCAAAATTAACCTCGGATTTGGCGATCCCTGAAATTTGGTGTCTATAAAGATGAGACG 249
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaAlaArgThrAlaGlnThr 100
DB 250 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTCTCGCAGGCGACAGACT 309
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 310 CCGGGGGGACCTGGCGCAGCTACGCGTGGCTGGCGATTTCTCGCAAAATTAACAGCGTT 369
QY 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
DB 370 AAGCGTGTGGGTGAGAACCCACCGACCTGGCCGACCATTAAGACGCTCTTAACCTCGCA 429
QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 430 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTGACTTCGAT 489
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 490 GCACCTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGAGTAGTGTCTTCCATGGCTGC 549
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
DB 550 TGCATATACCCCAACCGGTATCGACCTTACGCTGGAACAAATGGCAACACTGGCACAACTC 609
QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
DB 610 TCCGTTGAGAAAGCTGGTTACCGCTGTTGACTTCGCTTACAGGGTTTGGCCGTGTGT 669
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleVal 240
DB 670 CTGGAAGAAGATGCTCAAGAGACTGCGCGCTTTTCGCGCTATGCATAAAGAGCTGATTGTT 729
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
DB 730 GCAGTTCCTACTCTAAACATTTGGCCCTGTACACAGCGGTGTTGGCGCTGTGTACTCTG 789
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
DB 790 GTTGTCTGCCAGAGTAACCGTGTGATCGGCAATTCAGCCAAATGAAGCGGCGATTCGC 849
QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
DB 850 GCTAACTACTCTAACCCACGACACACGCGGCTTCTGTGTTGGCCACCATCTCTGAGCAAC 909
```

```
QY 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
DB 910 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGGTATTCAGCGT 969
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
DB 970 ATGCGTCAAGTGTTCGTCAATACGCTGCAGGAAAGGCGCAACCGCAGCTTCAGCTTT 1029
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
DB 1030 ATCATCAACAGAACGCGATTTCTCTTCAGTGGCTCGACAAAGAACAGTCTGCGT 1089
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 1090 CTGCGCGAGAGTTCGCGTATATGCGTTCGTTCTGCTCGCTAAATGTGGCGGAGTG 1149
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
DB 1150 ACACCAGATAACATGCTCGCTGCGAAGCGATTGTGGCAGTGTCTG 1197

RESULT 3
AR095441
LOCUS AR095441 1331 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 23 from patent US 6004773.
ACCESSION AR095441
VERSION AR095441.1 GI:10023318
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1331)
AUTHORS Araki, M., Sugimoto, M., Yoshihara, Y. and Nakamatsu, T.
TITLE Method for producing L-lysine
JOURNAL Patent: US 6004773-A 23 21-DEC-1999;
FEATURES
source
1..1331
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 8.08e-149 Length: 1331
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x AR095441 (1-1331)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 10 ATGTTTGAGAACATTACCGCGCTCTCGGACCGCATTCGGCCCTGGCCGATCTGTTT 69
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
DB 70 CGTGGCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 129
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
DB 130 GGCACAAACCCCGCTACTGACCGCTGAAAGAGGCTGAACAGTATCTGCTCGGCGATG 189
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 190 ACCACCAAAATTAACCTCGGATTTGGCGATCCCTGAAATTTGGTGTCTATAAAGATGAGACG 249
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaAlaArgThrAlaGlnThr 100
DB 250 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGCTCTCGCAGGCGACAGACT 309
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
```

Db 310 CCGGGGGGACCTGGCGCAGTACGCGTGGCTGGCGATTTCTCGCAAAAATACCAGCGTT 369

Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140

Db 370 AAGCGTGTGGTGAGCAACCAAGCTGGCGCAACCAATAAGAGCGCTTTAACTCTGCA 429

Qy 141 GlyLeuGluValArgGluTrpAlaTyrTrpAseAlaGluAenHisThrLeuAsePheAse 160

Db 430 GGTCTGGAAGTTCGTGAATACCGCTTATTATGATGGGAAAATCACACTCTTACATTCGAT 489

Qy 161 AlaLeuLeuAseSerLeuAseGluAlaGlnAlaGlyAseValValLeuPheHisGlyCys 180

Db 490 GCACCTGATTACAGCTGTAATGAGCTAGGCTGGGACGTAGTGTCTCCATGGCTGC 549

Qy 181 CysHisAseProThrGlyLeuAseProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200

Db 550 TGCCATAACCCACCGGTATCGACCTACGCTGGAACAATGGCAACACTGGCAACAATC 609

Qy 201 SerValGluLysGlyTrpLeuProLeuPheAsePheAlaTyrGlnGlyPheAlaArgGly 220

Db 610 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACAGGGTTTTGCCCGTGT 669

Qy 221 LeuGluGluAseAlaGluGlyLeuArgAlaPheAlaAseMetHisLysGluLeuVal 240

Db 670 CTGGAAAGAGTGTGAAGGACTGGCGGCTTTCCGGGCTATGCATAAAGAGCTGATTGT 729

Qy 241 AlaSerSerTrpSerLysAenPheGlyLeuTrpAseGluArgValGlyAlaCysThrLeu 260

Db 730 GCCAGTTCTACTCTAAAACCTTGGCGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 789

Qy 261 ValAlaAseSerGluThrValAseArgAlaPheSerGlnMetLysAlaAlaLeuArg 280

Db 790 GTTGCTGGCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCGC 849

Qy 281 AlaAseTrpSerAseProAlaHisGlyAlaSerValValAlaThrileLeuSerAen 300

Db 850 GCTAACTACTCTAACCCACAGCACACGGGCTTCTGTTGGCCACCATCTGAGCAAC 909

Qy 301 AseAlaLeuArgAlaLeuTrpGluGlnGluLeuThrAseMetArgGlnArgileGlnArg 320

Db 910 GATGCGTACGTGGCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTACGCGT 969

Qy 321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAseAseArgAsePheSerPhe 340

Db 970 ATGCGTCAGTTGTTCGTCAATACGCTGCAGGAAAAGGCGCAACCGCGACTTCAGCTTT 1029

Qy 341 IleIleLysGluAseGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

Db 1030 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTGACCAAGAACAGTGTGCGT 1089

Qy 361 LeuArgGluGluPheGlyValTrpAlaValAlaSerGlyArgValAseValAlaGlyMet 380

Db 1090 CTGGCGAAGAGTTTGGCGTATATGCGTTCGTTCTGCTGGCTGAATATGGCGGGATG 1149

Qy 381 ThrProAseAseMetAlaProLeuCysGluAlaIleValAlaValLeu 396

Db 1150 ACACAGATAACATGGCTCGCTGTCGGAAGCGATTGTGGCAGTGCTG 1197

RESULT 4

E16758 LOCUS E16758 1331 bp DNA linear PAT 28-JUL-1999

DEFINITION gDNA encoding aspartate aminotransferase (AAT).

ACCESSION E16758

VERSION E16758.1 GI:5711441

KEYWORDS JP 1998215883-A/16.

SOURCE Escherichia coli

ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 1331)

AUTHORS Araki, M., Sugimoto, M., Yoshihara, Y. and Nakamatsu, W.

TITLE PRODUCTION OF L-LYSINE

JOURNAL Patent: JP 1998215883-A 16 18-AUG-1998;

COMMENT

AJINOMOTO CO INC

OS Escherichia coli

PN JP 1998215883-A/16

PD 18-AUG-1998

PF 03-DEC-1997 JP 199733238

PR 05-DEC-1996 JP 96P 325659

PI ARAKI MASAYUKI, SUGIMOTO MASAKAZU, YOSHIHARA YASUHIKO, PI

NAKAMATSU WATARU

PC C12N15/09, C12N1/21, C12P13/08, (C12N15/09, C12R1:15), (C12N1/21, C12R1:15),

PC (C12P13/08, C12R1:15);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC Location/Qualifiers

Key

FT source 1..1331

FT /organism='Escherichia coli'

FT /strain='JM109'

FT 5'UTR 1..9

FT CDS 10..1197

FT /product='aspartate aminotransferase' FT

FT /gene='aspC'

FT 3'UTR 1198..1331.

FEATURES

source

1..1331

/organism='Escherichia coli'

/mol_type='genomic DNA'

/db_xref='taxon:562'

ORIGIN

Alignment Scores:

Pred. No.: 8,08e-149 Length: 1331

Score: 2045.00 Matches: 396

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x E16758 (1-1331)

Qy 1 MetPheGluAseNileThrAlaAlaProAlaAseProIleLeuGlyLeuAlaAseLeuPhe 20

Db 10 ATGTTTGAACAATTACCGCCGCTCTCGCGACCCGATTCGGGCTCGCCCATCTGTTT 69

Qy 21 ArgAlaAseGluArgProGlyLysIleAseLeuGlyIleGlyValTrpLysAseGluThr 40

Db 70 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGCG 129

Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTrpLeuLeuAseGlu 60

Db 130 GGCAAAACCCCGTACTGACCAGCGTGAAAAGGCTGAACAGTATCTGCTCGCAAAATGAA 189

Qy 61 ThrThrLysAseNileLeuGlyIleAseGlyIleProGluPheGlyArgCysThrGlnGlu 80

Db 190 ACCACCAAAATTAACCTCGGCATTCGCGCATCCCTGAATTTGGTTCGCTGACAGGAA 249

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAseAseLysArgAlaArgThrAlaGlnThr 100

Db 250 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGCGCACGACAGACT 309

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAsePheLeuAlaLysAseThrSerVal 120

Db 310 CCGGGGGGCACTGGCGGCATACGCGTGGCTGGCGGATTTCTCGGCAAAAATACCAGCGT 369

Qy 121 LysArgValTrpValSerAseProSerTrpProAseNileLysSerValPheAseSerAla 140

Db 370 AAGCGTGTGGGTGAGCAACCCAGCTGGCGAACCATTAAGAGCGTCTTTAACTCTGCA 429

Qy 141 GlyLeuGluValArgGluTrpAlaTyrTrpAseAlaGluAseNileThrLeuAsePheAse 160

Db 430 GGTCTGGAAGTTCGTGAATACCGCTTATTATGATGGGAAAATCACACTCTTGAATTCGAT 489

```
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
|
|
|
Db 490 GCACCTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGCTGC 549
|
|
|
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
|
|
|
Db 550 TGCATAAACCCACCGGTATCGACCTCGAACAATGGCAACACACTGGCACAACTC 609
|
|
|
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
|
|
|
Db 610 TCGTTTGAGAAAGCTGTGTTACCGCTGTTTGACTTTCCTTACCAAGGTTTGGCCGTGGT 669
|
|
|
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleVal 240
|
|
|
Db 670 CTGGAAGAGAGCTCTGAAGGACTCGCGCTTTCGCGCTATGCTAATAAGAGCTGATTGTT 729
|
|
|
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
|
|
|
Db 730 GCCAGTTCTTACTCTATAAATCTTGGCTGTACCAAGCGGTGTGGCGCTTGACTCTG 789
|
|
|
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
|
|
|
Db 790 GTTGTCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGCATTCGC 849
|
|
|
QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
|
|
|
Db 850 GCTAACTACTCTAAACCCACGACGACGCGGCTTCTGTGTTGCCACCATCTCTGAGCAAC 909
|
|
|
QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
|
|
|
Db 910 GATCGCTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCGCGTATTCAGCGT 969
|
|
|
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
|
|
|
Db 970 ATGCGTCAGTTGTCGTAATACGCTGCAGGAAAAAGCGCAACCGCGCACTTCAGCTTT 1029
|
|
|
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
|
|
|
Db 1030 ATCATCAACAGACAGCGCATGTTCTCTCAGTGGCTGCACAAAGAACAGTGTGCGT 1089
|
|
|
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
|
|
|
Db 1090 CTGCGGAAAGAGTTTGGCGTATATGCGTTGCTTCTGTCGCTAAATGTGGCGGATG 1149
|
|
|
QY 381 ThrProAsnMetAlaProLeuGluAlaIleValAlaValLeu 396
|
|
|
Db 1150 ACACCAAGATAACATGCTCGCTGTGCGAAGCGATTGTGCGCAGTGTGCTG 1197
|
|
|
RESULT 5
ECASPC 1415 bp DNA linear BCT 12-SEP-1993
LOCUS E. coli aspC gene for aspartate aminotransferase.
DEFINITION X03629
ACCESSION X03629.1 GI:41010
VERSION X03629.1
KEYWORDS aminotransferase.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
REFERENCE 1 (bases 1 to 1415)
AUTHORS Hunter,N.G., Finlay,M.E., Primrose,S.B., Parker,D.M. and
Edwards,R.M.
TITLE The cloning and sequence analysis of the aspC and tyrB genes from
Escherichia coli K12. Comparison of the primary structures of the
aspartate aminotransferase and aromatic aminotransferase of E. coli
with those of the pig aspartate aminotransferase isoenzymes
JOURNAL Biochem. J. 234 (3), 593-604 (1986)
PUBMED 3521591
COMMENT Data kindly reviewed (25-JUN-1986) by M. Hunter.
FEATURES
Location/Qualifiers
1..1415 source
```

```
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
68..73
/notes="pot. promoter signal"
91..96
/notes="pot. promoter signal"
104
/notes="mRNA start site"
127..130
/notes="ribosome binding site"
138..138
/notes="unnamed protein product; aspartate aminotransferase
(aspC)"
/codon_start=1
/transl_table=11
/protein_id="CAA27279.1"
/db_xref="GI:41011"
/db_xref="GOA:P00509"
/db_xref="UniProt/Swiss-Prot:P00509"
/translation="MFENITAAPADPILGLADLFRADRPGRKINLIGVVKDETGKTP
VLTSVKKAQVILLENETKNVIGIDGIPFEGCTOELLFGKGSALINDKRAFTAQTPG
GTGARLRVADFLAKNTSVKRVVSNPSWPHKSVENSAGLEVREYAYDAENHTLDFD
ALINSENEAQAQGVVLFHGCCHNPIDTLEQWQTLAQLSEKGVLPDFDPAQGFQFA
RGLSEDAEGLRAFAHKELELIVASSYSNFGLYNERVGAETLVAADSETVDPAFSQMK
AAIRANYSNPPAHGASVVAITILNDALRAIWEQELTDMQRQIQRMQLFVNTLQEKGA
NRDFSIIKQNGMFSFSLTKQVLRLEEFQVYVAVSGRVNVAGMTPDNMAPLCEAI
VAVL"
Alignment Scores:
Pred. No.: 8,69e-149 Length: 1415
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0
US-10-673-786A-2 (1-396) x ECASPC (1-1415)
QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
|
|
|
Db 138 ATGTTTGAGAACATTAACCGCGCTCTCGCCGACCCCGATTCTGGGCGCTGCGCATCTGTTT 197
|
|
|
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
|
|
|
Db 198 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGGAATTGGTGTCTATAAAGATGAGCG 257
|
|
|
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
|
|
|
Db 258 GGCACAAACCCCGGTACTGACACGCGTGAANAAGGCTGAACAGTATCTGCTCGAANAATGAA 317
|
|
|
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
|
|
|
Db 318 ACCACCAAAATATACCTCGCATTCAGCGCATCCCTGGAATTTGGTCTGCTGCACTCAGGAA 377
|
|
|
QY 81 LeuLeuPheGlyLysGlySerAlaIleAsnAspLysArgAlaArgThrAlaGlnThr 100
|
|
|
Db 378 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGACGCGACAGACT 437
|
|
|
QY 101 ProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
|
|
|
Db 438 CCGGGGGGACCTGGCGCACTACGCTGGCTGCGGATTTCTTGGCAAAAATAATACCGGTT 497
|
|
|
QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
|
|
|
Db 498 AAGCGTGTGTGGTGGAGCAACCAAGCTGGCGCAACCAATAGAGCGCTCTTTAACTCTGCA 557
|
|
|
QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
|
|
|
Db 558 GGTCTGGAAGTTTCGTGAATACGCTTATTATGATGCGGAAAAATCACACTCTTGAATTCGAT 617
|
|
|
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
|
|
|
```

```

Db      618  GCACCTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTCTCCATGCTGC 677
Qy      181  CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db      678  TGCCTAAACCCACCGGTATCGACCTAGCTGGAAACATGCAACACATGGCACACATC 737
Qy      201  SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db      738  TCCGTTTGAGAAAGCGTGTATACCGCTGTGTGACTTCGCTTACAGGGTTTGGCCGCTGT 797
Qy      221  LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLeuVal 240
Db      798  CTGGAAGAAGATGCTGAAGACTGGCGCTTTGCGGGCTATGCTAAAGAGCTGATTGTT 857
Qy      241  AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db      858  GCCAGTTCTCTCTTAAACCTTTGGCTGTACACAGCGGTGTGGCGCTTGTACTCTG 917
Qy      261  ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaLeuArg 280
Db      918  GTTGTGCGCACAGTGAACCGTTGATCGGCATTGAGCCAAATGAAGCGCGATTGCG 977
Qy      281  AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db      978  GCTAACTACTCTAAACCCACGACACACGCGCTTCTGTGTTGCCACCATCTCTGAGCAAC 1037
Qy      301  AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db      1038  GATGCGTTACGTCGATTTGGGAACAAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 1097
Qy      321  MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db      1098  ATGCGTCAGTTGTCGTAATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1157
Qy      341  IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db      1158  ATCATCAAAACAGAACGCGATGTTCTCTTCAGTGGCGCTGACAAAGAACAAAGTGTGCGT 1217
Qy      361  LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db      1218  CTGCGCGAAGAGTTTGGCGTATATGCGTTGCTTCTGTCGCGTAAATGTGGCGGGATG 1277
Qy      381  ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db      1278  ACACACAGATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTGTCTG 1325

```

```

RESULT 6
LOCUS   108484
DEFINITION Sequence 9 from Patent WO 8700202.
ACCESSION 108484
VERSION 108484.1 GI:588804
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1468)
AUTHORS Edwards,M.R., Taylor,P.P., Hunter,M.G. and Fotheringham,I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 9 15-JAN-1987;
FEATURES Location/Qualifiers
         1..1468
            /organism="unknown"
            /mol_type="unassigned DNA"

```

ORIGIN

Alignment Scores:

Pred. No.:	9, 07e-149	Length:	1468
Score:	2045.00	Matches:	396
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:		Indels:	0

```

DB:
US-10-673-786A-2 (1-396) x 108484 (1-1468)
Gaps: 0
Qy      1  MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db      191  ATGTTTGAACAATACCGCGCTCTCGCCGACCCGATTCTGGGCTTGGCCGATCTGTTT 250
Qy      21  ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db      251  CGTCCGATGAACGTCGCGCAAAATTAACTCCGGATTGGTGTCTATATAAGATGAGACG 310
Qy      41  GlyLysThrProValLeuThrSerValLysIleAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db      311  GGCAAAACCCCGGTACTACCGAGCGTGAAGGCTGAACAGTATCTGCTCGAAATGAA 370
Qy      61  ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db      371  ACCACCAAAATACCTCGGCATTGACCGCATCCCTGAAATTTGGTGGCTGCCTCAGGAA 430
Qy      81  LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db      431  CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCCTCGCACGCACAGCT 490
Qy      101  ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db      491  CCGGGGGCACCTGGCGGCACCTACGCGTGGCTGCCGATTTCTGGCAAAAATACCAAGCT 550
Qy      121  LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db      551  AAGCGTGTGGTGAGCAACCCAGCTGGCGCAACCATAGAGCGCTCTTAACTCTGCA 610
Qy      141  GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db      611  GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAATTCACACTCTTGACTTCGAT 670
Qy      161  AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db      671  GCACGTGATTAAACAGCTGAATGAAGCTCAGCGCTGGCGACGTAGTGTCTGTTCCATGGCTGC 730
Qy      181  CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db      731  TGCCATAACCCACACCGGTATCGACCTAGCTGGAAACATGGCAAAACACTGGCACACATC 790
Qy      201  SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db      791  TCCGTTGAGAAAGCGTGTGTTACCGCTTTGACTTCGCTTACCAAGGCTTTTCCCGTGTG 850
Qy      221  LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuLeuVal 240
Db      851  CTGGAAGAAGATGCTGAAGGACTGGCGCTTTTCCGCGCTATGCATAAAGAGCTGATTGTT 910
Qy      241  AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db      911  GCCAGTTCTCTACTCTAAAAACCTTTGGCCTGTACACAGAGCGGTGTGGCGCTGTACTCTG 970
Qy      261  ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaLeuArg 280
Db      971  GTTGTGCGGACAGTGAACCGTTGATCGCGCATTCAGCAATGAAAGCGCGGATTCGC 1030
Qy      281  AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db      1031  GCTAACTACTCTAAACCCACGACACACGCGCTTCTGTTGTTGGCACCCTTCCTGAGCAAC 1090
Qy      301  AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db      1091  GATGCGTTACGTCGCGATTGGAACAAAGAGCTGACTGATATGCGCCAGCGTATTACGCT 1150
Qy      321  MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe 340
Db      1151  ATGCGTCAGTTGTTGTCGTAATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1210
Qy      341  IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

```

Db 1211 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTGCACAAAGAACAAAGTGGCGT 1270
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1271 CTGGCGGAAGAGTTTGGCGTATATGCGTGTCTCTGTCGCGTAATGTGGCGGGATG 1330
 QY 381 ThrProAspAsnMetAlaProLeuGluGluAlaValAlaValLeu 396
 Db 1331 ACACCAAGATAACATGCTCGCTGTGCGAAGCGATTGTGGCAGTGTG 1378

RESULT 7

LOCUS I08487 3659 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 15 from Patent WO 8700202.

ACCESSION I08487

VERSION I08487.1 GI:588807

KEYWORDS Unknown.

SOURCE Unclassified.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3659)

AUTHORS Edwards, M.R., Taylor, P.P., Hunter, M.G. and Fotheringham, I.G.

TITLE COMPOSITE PHARMIDS FOR AMINO ACID SYNTHESIS

JOURNAL Patent: WO 8700202-A 15 JAN-1987;

FEATURES Location/Qualifiers

1..3659 /organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores: 2.68e-148 Length: 3659

Pred. No.: 2045.00 Matches: 396

Score: 2045.00 Conservative: 0

Percent Similarity: 100.0% Mismatches: 0

Best Local Similarity: 100.0% Indels: 0

Query Match: 100.0% Gaps: 0

DB: 6

US-10-673-786A-2 (1-396) x I08487 (1-3659)

QY 1 MetPheGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20

Db 2344 ATGTTTGAACAATACCGCGCTCTGCGCAGCCCGATTCTGGCGCTGGCGCATCTGTTT 2403

QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleValTyrLysAspGluThr 40

Db 2404 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGCG 2463

QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60

Db 2464 GGCNAACCCGGTACTGACCGCTGAANAAGCTGAACATGATCTGTCGNAATGAA 2523

QY 61 ThrThrLysAsnTyrLeuGlyLeuAspGlyLeuProGluPheGlyArgCysThrGlnGlu 80

Db 2524 ACCACCAAAATTTACCTCGCATGACCGCATCCCTGAAATTTGGTGGCTGCACCTAGGAA 2583

QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100

Db 2584 CTGCTGTTTGGTAAAGTAGCGCTTCATGACAAACGCTGCTCGCACGACAGCT 2643

QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120

Db 2644 CCGGGGGGACCTGGCGCACTACCGTGGCTGCGGATTTCTGGCAAAAATACACCGTT 2703

QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140

Db 2704 AAGCGTGTGGTGAGCAACCAACGCTGCCGCAACCAATAGCGCTCTTAACCTGCA 2763

QY 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160

Db 2764 GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAAATCACACTCTTCACTTCGAT 2823

QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180

Db 2824 GCATGATTAACAGCGCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCATGCTGC 2883

QY 181 CysHisAsnProThrThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200

Db 2884 TCCCATTAACCCAAACCGTATCGACCTACCTGGAAACAATGGCAAACTGGGCACAATC 2943

QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220

Db 2944 TCCGTTTGAGAAAGCGTGGTTACCGCTGTTTGCCTTACCAGGTTTGGCCCGTGGT 3003

QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240

Db 3004 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCATTAAGAGCTGATGTT 3063

QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260

Db 3064 GCCAGTTCTCTCTTAAACCTTTGGCTGTACACAGCGGTGTGGCGCTGTGTACTCTG 3123

QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaIleArg 280

Db 3124 GTTGTGCGCAGCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTGCG 3183

QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValAlaThrLeuSerAsn 300

Db 3184 GCTAACTACTCTAAACCCACAGCAGCGCGCTTCTGTGTTGCCACCATCTCTGAGCAAC 3243

QY 301 AspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320

Db 3244 GATCGGTACGTGCGATTGCGGAACAAGAGCTGATGATGCGCCAGCGGTATTACGCGT 3303

QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAlaAsnArgAspPheSerPhe 340

Db 3304 ATGCGTCAGTTGTTGTCATACGCTGCGAGAAAGGCGCAACCGCAGCTTCAGCTTT 3363

QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

Db 3364 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCTGCACAAAGAACAAAGTGTGCGT 3423

QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380

Db 3424 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTGGCCGGATG 3483

QY 381 ThrProAspAsnMetAlaProLeuGluAlaIleValAlaValLeu 396

Db 3484 ACACCAAGATAACATGCTCGCTGCGAAGCGATTGTGGCAGTGTGCTG 3531

RESULT 8

D90731/c 11470 bp DNA linear BCT 25-DEC-2002

LOCUS Escherichia coli K12 genomic DNA. (21.2 - 21.5 min).

DEFINITION D90731 AB001340

ACCESSION D90731.1 GI:1651452

VERSION Complete and shotgun sequencing; ycbB; aspC; ompF; tolF; cmlB; coa;

KEYWORDS cry; asnS; tss; pncB; pepN; ycbE; yzeB.

SOURCE Escherichia coli K12

ORGANISM Escherichia coli K12

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

REFERENCE 1

AUTHORS Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,

Ikenoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K.,

Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T.,

Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishimoto, H.,

Nishio, Y., Saito, N., Sampa, G., Seki, Y., Tagami, H., Takemoto, K.,

Wada, C., Yamamoto, Y., Yano, M. and Horiuchi, T.

A 718-kb DNA sequence of the Escherichia coli K-12 genome

corresponding to the 12.7-28.0 min region on the linkage map

DNA Res. 3 (3), 137-155 (1996)

PUBLISHED 8905232

REFERENCE 2

AUTHORS Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T.,

Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y., and Yano, M.

The systematic sequencing of the *Escherichia coli* genome in Japan
Unpublished
3 (bases 1 to 11470)

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Direct Submission
Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan
(E-mail: hmori@gtc.aist-nara.ac.jp, Tel: 81-7437-2-5660, Fax: 81-7437-2-5669)

COMMENT

Collaboration Information:

Project:
The Japan E.coli genome DNA sequencing project

Group:
The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)
Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.

Headed by:

Name: Takashi Horiuchi
Address: National Institute of Basic Biology, Okazaki, 444, Japan
E-mail: kishorienibb.ac.jp

Information operator:

Name: Hirotsada Mori
Address: NARA Institute of Science and Technology,
Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL:

The Japan E. coli genome database
http://bsw3.aist-nara.ac.jp.

FEATURES source

Location/Qualifiers
1..11470
/organism="Escherichia coli K12"
/mol_type="genomic DNA"
/strain="K12"
/db_xref="taxon:83333"
/map="21.2-21.5 min"
/clone="Kohara clone #219"
/note="Nucleotide position 983395-994864 from the initiation site of *thrA* (0 min.).
This clone is from Kohara lambda miniset library."

gene

1..47
/gene="ycbB"

-1..47

CDS

/gene="ycbB"
/note="ORF_ID:0218#11
similar to PIR Accession Number PH0205"
/codon_start=3
/transl_table=11
/product="Hypothetical protein 63 (MukB 3' region)"
/protein_id="BAA35677.1"
/db_xref="GI:4062495"
/translation="SSSQIVSKAEQLIR"
228..776

CDS

/note="ORF_ID:0219#1
similar to PIR Accession Number F64039"
/codon_start=1
/transl_table=11
/product="Hypothetical protein H11666"
/protein_id="BAA35678.1"
/db_xref="GI:4062496"

/translation="MDKFEDANRRKLLALGGVALGAAILPTPAFATLSTPRPRLTLNN
LHTESKAEFFDGRGTQBELAKLNHFFRDYRANKIKSIDPGLFDQIYRQGLGTR
KPVLIIGYSRIDTNNSRAKRSVGAKKSYHTKGQAMDHFHIEGIALSNIRKAALSRA
GGVGYPRSNFVHIDTGPARHW"

CDS

803..11450
/note="ORF_ID:0219#2
similar to SwissProt Accession Number P54501"
/codon_start=1
/transl_table=11
/product="Hypothetical 23.2 kd protein in soda-comGA
intergenic region."
/protein_id="BAA35679.1"
/db_xref="GI:4062497"
/translation="MNYRIIPVTAFSONCSLIWEQRTLAALVDPGGDAEKIKQEVDD
SGLTMIQILTHGLHDHVAALAQHYGVFVGPBEKEDEFWLQGLPAQSRMFLGEEC
QPLTPDRLNEGDTISICNTVLQVLCHPGTHFVGFDDRAKLLISGDIVFKGGVGR
SDPRGDHNLISIKKLLPLGDDVIFPGHGLSTLGYERLHNPFQDEMPVW"
complement(1672..2862)
/gene="aspC"
complement(1672..2862)
/gene="aspC"
/note="ORF_ID:0219#3
similar to PIR Accession Number A00598"
/codon_start=1
/transl_table=11
/product="Aspartate transaminase (EC 2.6.1.1)"
/protein_id="BAA35680.1"
/db_xref="GI:1651453"
/translation="MFSNITAAPADPILGLADLFRADPRGKINLIGVYKDETKTP
VLTSVKAEQYLLENETKYNLVDGIPFCFQCTQELLFGKGSALLINDKRACTATPG
GTGALRVAAFDLAKNTSVKRVNSNPNHKSVFNSAGLEVREYATYDAENHTLDFG
ALINSLAEQAQDVVLFHGCNPTGIDPTLEQWTLAQLSVKGMFLPFDFAFQGFA
RLGEDAEGLRAFAAMHKLIVASSYKNGFLYNERVGACTLVADSETVDFRAFQOMK
AAIANTYNNPAGASVVATILSNDLRALIWEQSLTDMRQIRMRQLFVNTLQSKGA
NRDEFILKQNGMFSFSLTKQVLRUREBEGVAVASGRVNVAGTMTFDNMAPLCEAI
VAVL"
complement(3047..4135)
/gene="cm1B"
/note="synonyms: coa, cry, ompF, tofF"
complement(3047..4135)
/gene="cm1B"
/note="ORF_ID:0219#4
similar to SwissProt Accession Number P02931"
/codon_start=1
/transl_table=11
/product="Outer membrane protein f precursor (outer
membrane protein la, la, or b)."
/protein_id="BAA35681.1"
/db_xref="GI:1651454"
/translation="MKRNILAVIPALLVAGTANAAEIKNGDKNKVDLYKAVGLHY
PSKNGNSYGGNGDMTYARLFGKETQINSDLTCYGQWYFNQGNSEGDADQGNK
TRLAPAGLKYADVGSFDYGRNYGVYDALGYTDLMPFGGTATYDDDFVGRVGGVAT
YRNSFGLVDGLNFAVQYLKNERDTRRRSNGDVGGSISYEYEGFVIGVGAADR
TNLQEAQPLNGKKAQWATGLKYDANNIYLAANYGNRNPATPTNKNTNSGFANKT
QDVLLVAQYQDFGLRPSIATYKSKADVEGIGVDLVNFEVGATYTFNKNMSTVVD
YIINQIDSNKLVGSDDTVAVGIVYQF"

CDS

/gene="aasS"
complement(4738..6138)
/gene="aasS"
/note="synonym: tes"
complement(4738..6138)
/gene="aasS"
/note="ORF_ID:0219#5
similar to SwissProt Accession Number P17242"
/codon_start=1
/transl_table=11
/product="Asparaginyl-tRNA synthetase (EC 6.1.1.22)
(asparagine-tRNA ligase) (aasRS)."
/protein_id="BAA35682.1"
/db_xref="GI:1651455"
/translation="MSVVPVADVLRQGRVAVDSEVTVRGVRTRRDSKAGISFLAVYDG
SCFPVQAVINNSLPYNYEDVRLTTCGSLVTGKVVASPGQGFQFEIQAQSKVEVAGW
VEDPDTYPMAAKRHSIEYLRVAHLRPRTNLIGAVRVHRLAQLHRRFPQGFVW
STPLITASDTGEGEMFRVSTLDLENLPRNDQGRVDFDKDFFGKESFLTVSQNLNET

CDS

gene
complement(3047..4135)
/gene="cm1B"
/note="synonyms: coa, cry, ompF, tofF"
complement(3047..4135)
/gene="cm1B"
/note="ORF_ID:0219#4
similar to SwissProt Accession Number P02931"
/codon_start=1
/transl_table=11
/product="Outer membrane protein f precursor (outer
membrane protein la, la, or b)."
/protein_id="BAA35681.1"
/db_xref="GI:1651454"
/translation="MKRNILAVIPALLVAGTANAAEIKNGDKNKVDLYKAVGLHY
PSKNGNSYGGNGDMTYARLFGKETQINSDLTCYGQWYFNQGNSEGDADQGNK
TRLAPAGLKYADVGSFDYGRNYGVYDALGYTDLMPFGGTATYDDDFVGRVGGVAT
YRNSFGLVDGLNFAVQYLKNERDTRRRSNGDVGGSISYEYEGFVIGVGAADR
TNLQEAQPLNGKKAQWATGLKYDANNIYLAANYGNRNPATPTNKNTNSGFANKT
QDVLLVAQYQDFGLRPSIATYKSKADVEGIGVDLVNFEVGATYTFNKNMSTVVD
YIINQIDSNKLVGSDDTVAVGIVYQF"

CDS

gene
complement(4738..6138)
/gene="aasS"
/note="synonym: tes"
complement(4738..6138)
/gene="aasS"
/note="ORF_ID:0219#5
similar to SwissProt Accession Number P17242"
/codon_start=1
/transl_table=11
/product="Asparaginyl-tRNA synthetase (EC 6.1.1.22)
(asparagine-tRNA ligase) (aasRS)."
/protein_id="BAA35682.1"
/db_xref="GI:1651455"
/translation="MSVVPVADVLRQGRVAVDSEVTVRGVRTRRDSKAGISFLAVYDG
SCFPVQAVINNSLPYNYEDVRLTTCGSLVTGKVVASPGQGFQFEIQAQSKVEVAGW
VEDPDTYPMAAKRHSIEYLRVAHLRPRTNLIGAVRVHRLAQLHRRFPQGFVW
STPLITASDTGEGEMFRVSTLDLENLPRNDQGRVDFDKDFFGKESFLTVSQNLNET

YACALSKIYTFGTFRANSENSTRHIAEFWMLEPEVAFANLNDIAGLAEMLYKVFKA
VLEERADDMKFAERVDKAVSRLEFIEADFAQVDYTDVITLENCRGRFPENPVWG
VLSSEHYLAEEHFKAQVAVKQYKDKAFYMRLENDKTKVAAMDVLAPGIBIIG
GSDRELDVDRLEMLKNEKEDYWMYRDLRRYGVTHVPSGFGLEFRLIAVVTGVON
VRDVIPTPTBNASF"

complement (6307..7509)

/gene="pncB"

complement (6307..7509)

/gene="pncB"

/note="ORF ID:0220#1"

similar to PIR Accession Number JQ0756"

/codon_start=1

/transl_table=11

/product="Nicotinate phosphoribosyltransferase (BC

2.4.2.11)"

/protein_id="BAA35683.1"

/db_xref="GI:1651456"

/translation="MTQFASPVHLSDTDAYKLHQQAQVHHYVDVHVAABFRCRGD
DLGIYADAIHQVQAMHRLQDDEYQWLSALPFKADYLNWLEFRNPEQVTSN
DNGKLDIRLSQPMREVLWEVPLLAIVISGMHRYSPQADVAQALDTLESKLVDPSAL
TAGLDMSPHLDGFRRRFSREVOETIVKRLQDSFWFVGTSTNYDLARLSLTPMGTO
AHFWQAHQOIGSDPNSORAAALAAWLEYPDLGIALTDCITMDAFIRDGVEFASR
YQGLRHSDSDPVEWEKATAHYEKGIDPQSKTLVPSDNLDLRKAVELYRHFSSRVQL
SFQIGRLTCDIPQVKPLNIVIKLVECNKPKVAKUSDSPGKTIChDKKAFVRLRKAFD
LPHIKKAS"

gene 7775..10387

/gene="pncB"

7775..10387

/gene="pncB"

/note="ORF ID:0220#2"

similar to SwissProt Accession Number P04825"

Alignment Scores:

Pred. No.: 1,046-147 Length: 11470
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x D90731 (1-11470)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 2862 ATGTTTGAGAAATTACCGCCGCTCTCGCGACCCGATTCGGGCTGCGCCATCTGTTT 2803
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrIlyAspGluThr 40
Db 2802 CGTGGCGATGAACGTCGCCGCAAAATTAACCTCGGATTTGGTGTCTATAAGATGAGACG 2743
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 2742 GCGAAACCCCGGTACTACACGCGTGAAAGAGGCTGAACAGTATCTGCTCGAAATGAA 2683
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 2682 ACCACCAAAATATACCTCGGCAATGACGCGATCCCTGGAATTTGGTCGCTGCACTAGAA 2623
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 2622 CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGCTGTCGACGCGACAGACT 2563
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 2562 CCGGGGGGACATGGCGCACTACGCGGTGCGGATTTCTCGCAAAATAATACCGCGTT 2503
Qy 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
Db 2502 AAGCGTGTGGGTGAGCAACCAACAGCTGGCGCAACCAATGAGAGCGTCTTTAACTCTGCA 2443
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 2442 GGTCTGGAAGTTCGTGATACGCTTATATGATGCGGAAATACACACTCTTCACTTCGAT 2383

Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 2382 GCACTGATTAAACAGCGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGCTGC 2323
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrPheGlnThrLeuAlaGlnLeu 200
Db 2322 TGCCATTAACCCCAACCGGTATCGACCTACGCTGGAAACAATGGCAACACTGGCACAATC 2263
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaAspGly 220
Db 2262 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCCTCGCTTACAGGCTTTTGGCCGTGGT 2203
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 2202 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGGGCTATGCATAAAGAGCTGATTGTT 2143
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 2142 GCCAGTTCCTACTCTAAAACCTTGGCCTGTACAACGAGCGTGTGGCGCTTGTACTCTG 2083
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 2082 GTTGCTGCGCAGTGAACCGCTTGTATCGCGCTTTCAGTATGCGCCAAATGAAGCGGATTCGC 2023
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 2022 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 1963
Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1962 GATGCTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATGCCCGCAGGCTATTCAGCGT 1903
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe 340
Db 1902 ATGCGTCACTGTTGCTCAATACGCTGCGAGGAAAGGCGCAACCCGACATTCAGCTTT 1843
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1842 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAGCTGCTGCGT 1783
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1782 CTGCGCAAGAGCTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1723
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1722 ACACAGATAACATGCTCGCTGCGAAGCGATTTGTGCGAGTGTGCTG 1675

RESULT 9

AX370270/c

LOCUS

Sequence 87 from Patent WO0170776.

DEFINITION

AX370270

ACCESSION

AX370270.1

VERSION

AX370270.1

KEYWORDS

Escherichia coli

ORGANISM

Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

REFERENCE

1

Levy, S.B., Barbosa, T.M. and Alekshun, M.N.

Nimr compositions and their methods of use

Patent: WO 0170776-A 87 27-SEP-2001;

TRUSTEES OF TUFTS COLLEGE (US)

Location/Qualifiers

1. 14759

/organism="Escherichia coli"

/mol_type="unassigned DNA"

/db_xref="taxon:562"

ORIGIN

Alignment Scores:

Pred. No.: 1.41e-147 Length: 14759
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x AX370270 (1-14759)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 2755 ATGTTTGAGAACATTACCGCCCTCTGCCGACCCGATTCGGGCTCGCCGATCTGTT 2696
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 2695 CGTCCGATGAACGTCCTCCGCAAAATTAACTCTCGGATTTGGTGTCTATAAAGATGAGACG 2636
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 2635 GGCAAAACCCCGGTACTGACCAAGCGTGAAAAGGCTGAACAGTATCTGTCGAAAATGAA 2576
 Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 2575 ACCACAAAATTAACCTGGCATTGACGCGATCCCTGAATTTGGTGTGCTGCACTCAGGAA 2516
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 2515 CTGCTGTTTGGTAAAGTAGCGCTGATCAATGACAAACGTCGTCGACGACACAGACT 2456
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 2455 CGCGGGGCGACTGGCGCCTACGCGTGGCTGCGGATTTCTGGCAAAAAATACCAAGCGTT 2396
 Qy 121 LysArgValTTPValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 2395 AAGCGTGTGGGTGAGCAACCAAGCTGGCGCAACCAATGAGAGCGTCTTTAACTCTGCA 2336
 Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 2335 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATACACACTTTGACTTCGAT 2276
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2275 GCACGTGATTAACAGCGCTGAATGAAGCTCAGGCTGGCGAGTAGTGTCTGTTCCATGGCTGC 2216
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTTPGlnThrLeuAlaGlnLeu 200
 Db 2215 TGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCAACATC 2156
 Qy 201 SerValGluLysGlyTTPLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 2155 TCCGTTGAGAAAGCTGTTACCGCTGTTTGACTTCGCTTACCGGTTTGGCCGCTGGT 2096
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 2095 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTTCGGCGGTATGCATAAAGAGCTGATGTT 2036
 Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluAtrqValGlyValaCysThrLeu 260
 Db 2035 GCCAGTTCCTACTCTAAAACATTTGGCCTGTACACAGAGCGGTGGCGCTTGTACTCTG 1976
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 1975 GTTGTGCGGACAGTGAACCTGTTGATCGCGATTCAGCCAATGAAGCGCGATTCG 1916
 Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValAlaThrIleLeuSerAsn 300
 Db 1915 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTTGGCCACCATCTCTGAGCAAC 1856
 Qy 301 AspAlaLeuArgAlaIleTTPGlnGlnLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 1855 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCGCGGTATTCAGCGT 1796

Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPhe 340
 Db 1795 ATGGCTCAGTTGTTGCGCAATACGTCGAGAAAGGCGCAAAACCGGACTTTCAGCTTT 1736
 Qy 341 IleIleLysGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1735 ATCATCAACAGACGCGCATGTTCTCTTCACTGGCTGACAAAAGAACACAGTCTGCGT 1676
 Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1675 CTGCGCAAGAGTTTGGCGTATATCGGTTCTTCTGTCGCTAAATGTGCGCGGATG 1616
 Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1615 ACACAGATAACATGGCTCGCTGCGAAGCGATTGTGGCAGTCTG 1568
 RESULT 10
 D90730/c
 LOCUS
 DEFINITION
 D90730 AB001340
 ACCESSION
 VERSION
 D90730.1 GI:1651442
 COMPLETE and shotgun sequencing; msbA; ycaH; kdsB; ycbC; smtA;
 mukF; kicB; mukE; kicA; mukB; ycbB; aspC; ompF; tolF; cmlB; coa;
 cry; asnS; tss.
 SOURCE
 Escherichia coli K12
 ORGANISM
 Escherichia coli K12
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 REFERENCE
 1 Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K., Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y., Yano, M. and Horiuchi, T.
 A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map
 DNA Res. 3 (3), 137-155 (1996)
 2 Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.
 The systematic sequencing of the Escherichia coli genome in Japan
 Unpublished
 3 (bases 1 to 20604)
 Mori, H.
 Direct Submission
 Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan
 (E-mail: hmori@gtc.aist-nara.ac.jp, Tel: 81-7437-2-5660, Fax: 81-7437-2-5669)
 Collaboration Information:
 Project:
 The Japan E.coli genome DNA sequencing project
 Group:
 The Japan E.coli genome DNA sequencing group
 Members: (1995.4 - 1996.3)
 Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, Y., Oshima, T., Saito, N., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.
 Headed by:

Name: Takashi Horiuchi
 Address: National Institute of Basic Biology, Okazaki, 444, Japan
 E-mail: kishori@nibb.ac.jp
 Information operator:
 Name: Hirotada Mori
 Address: NARA Institute of Science and Technology,
 Ikoma, 630-01, Japan
 E-mail: hmori@nara.ac.jp
 URL:
 The Japan E. coli genome database
 http://bwa3.aist-nara.ac.jp
 http://bwa3.aist-nara.ac.jp

FEATURES

source

Location/Qualifiers
 1..20604
 /organism="Escherichia coli K12"
 /mol_type="genomic DNA"
 /strain="K12"
 /db_xref="taxon:83333"
 /map="20.9-21.3 min"
 /clone="Kohara clone #218"
 /notes="Nucleotide position 968194-988797 from the
 initiation site of ThrA (0 min.).
 This clone is from Kohara lambda miniset library."

gene
 1..723
 /gene="msbA"
 CDS
 <1..723
 /gene="msbA"
 /notes="ORF ID:0217#11
 similar to PIR Accession Number S27998"
 /codon_start=1
 /transl_table=11
 /product="MsbA protein."
 /protein_id="BAA35660.1"
 /db_xref="GI:1651443"
 /translation="EFRTVTFYPCGDDVPALRNINLKIPAGKTVALVGRSGSKSTIA
 SLITRFDIGEDILMDGDLREYTLASRNQVALVQNVHLFNDTVANNIAYATEQ
 YSREQIEEAARWAYAMDFINKMDGLDTVIGENGLLSGGQRIARALARDSPIL
 ILDEATSAIDTESRAIQALDELQKNTSLVIAHRLSTIEKADEIVVVEDGVIVERG
 THNDLLEHGVVLAQHKMQFGQ"
 720..1706
 /gene="ycaH"
 CDS
 720..1706
 /gene="ycaH"
 /notes="ORF ID:0218#1
 similar to PIR Accession Number S27999"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein E (msbA 3' region)"
 /protein_id="BAA35661.1"
 /db_xref="GI:4062487"
 /translation="MTEKTSGESPLRLLLPLSLWLYGLVSGAIRLCYKLLKRAWRA
 PVPVVGNTAGGKTPVWVWVLEQLOQGRGIRVGVSRGYGKKAESYPLLLSADTT
 TQAGDEPLIYQRTDAPVAVSPVRSDAKAILAQHPDQIIIVDDGLQYRLARDE
 IIVIDGVRFGNGLWLPAGPFRERAGRLKSDAVIVNGVPRSPRHPMLLPQAVNL
 RTGTCDAQLHEVAVMAGIHPPRFATLKMCGVQPKCVPLADHQSILNHDVLSALV
 SAGOTLVMTKQAKVACRAFEENWYLPVDAQLSGDEPAKLLTQLTLASGN"
 1743..2975
 /notes="unnamed protein product; ORF_ID:0218#2"
 /codon_start=1
 /transl_table=11
 /protein_id="BAA35662.1"
 /db_xref="GI:4062488"
 /translation="MSLPHLSADARNHLAAQGLNKKRRRASLEDIPATISRWSL
 QDITINIVARSPVLVFSRLGNYPAGWDESLARGELMEYWAHEACFMRSPRFLIRH
 RMLAPEKMGWYKQAMWQHEAEIAQLIHIHDKGVRSDAEHPHPRKSGAGWGWKPH
 KRLHGLEGTAGVWYTERNFQYDLTVMPDWDDEDLVQSQAEIIMLDNGARS
 LGTFRQWLADVYRLKRPALAAWREARAEQQQIIIVHVEKLGNLWLDHLLPLLERAL
 AKGLTATHSVSPDPVVDKRAEQLFDFSVRLFCYTPAPKROYGVFVPLLHRCQ
 LVGRMDAKHRTQGLEIVSLWLEQGIKPTTLQKGLRQAITDPANWQATRTVLTGCC
 POGFLDTCRTGWGIDPVA"
 3027..3209
 /notes="unnamed protein product; ORF_ID:0218#3"
 /codon_start=1
 /transl_table=11

gene
 720..1706
 /gene="ycaH"
 CDS
 720..1706
 /gene="ycaH"
 /notes="ORF ID:0218#1
 similar to PIR Accession Number S27999"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein E (msbA 3' region)"
 /protein_id="BAA35661.1"
 /db_xref="GI:4062487"
 /translation="MTEKTSGESPLRLLLPLSLWLYGLVSGAIRLCYKLLKRAWRA
 PVPVVGNTAGGKTPVWVWVLEQLOQGRGIRVGVSRGYGKKAESYPLLLSADTT
 TQAGDEPLIYQRTDAPVAVSPVRSDAKAILAQHPDQIIIVDDGLQYRLARDE
 IIVIDGVRFGNGLWLPAGPFRERAGRLKSDAVIVNGVPRSPRHPMLLPQAVNL
 RTGTCDAQLHEVAVMAGIHPPRFATLKMCGVQPKCVPLADHQSILNHDVLSALV
 SAGOTLVMTKQAKVACRAFEENWYLPVDAQLSGDEPAKLLTQLTLASGN"
 1743..2975
 /notes="unnamed protein product; ORF_ID:0218#2"
 /codon_start=1
 /transl_table=11
 /protein_id="BAA35662.1"
 /db_xref="GI:4062488"
 /translation="MSLPHLSADARNHLAAQGLNKKRRRASLEDIPATISRWSL
 QDITINIVARSPVLVFSRLGNYPAGWDESLARGELMEYWAHEACFMRSPRFLIRH
 RMLAPEKMGWYKQAMWQHEAEIAQLIHIHDKGVRSDAEHPHPRKSGAGWGWKPH
 KRLHGLEGTAGVWYTERNFQYDLTVMPDWDDEDLVQSQAEIIMLDNGARS
 LGTFRQWLADVYRLKRPALAAWREARAEQQQIIIVHVEKLGNLWLDHLLPLLERAL
 AKGLTATHSVSPDPVVDKRAEQLFDFSVRLFCYTPAPKROYGVFVPLLHRCQ
 LVGRMDAKHRTQGLEIVSLWLEQGIKPTTLQKGLRQAITDPANWQATRTVLTGCC
 POGFLDTCRTGWGIDPVA"
 3027..3209
 /notes="unnamed protein product; ORF_ID:0218#3"
 /codon_start=1
 /transl_table=11

gene
 CDS
 /protein_id="BAA35663.1"
 /db_xref="GI:4062489"
 /translation="MDHRLLEIIACPVCGKGLWYNQEKQELICKLDNLNLAFFLRDGPV
 LLETEARVLTADESKS"
 3206..3952
 /gene="kdsB"
 3206..3952
 /gene="kdsB"
 /note="ORF ID:0218#4
 similar to PIR Accession Number A26322"
 /codon_start=1
 /transl_table=11
 /product="3-deoxy-manno-octulosonate cytidylyltransferase
 (EC 2.7.7.38)"
 /protein_id="BAA35664.1"
 /db_xref="GI:1651444"
 /translation="MSFVVIIPARYASTRLPGKPLVDINGKPMIVHVLREARSGAER
 IIVATDHEDVARVAEAGEVCMTRADHQSQGLAEVVEKCAFSDDTVVNVQGDPEP
 MPATIIIRQVADNLQROQMATLAVPHINAEAEFNPVAKVVDLDEGVALYFSRATI
 WPRDRFAEGLETVDGNFLRHIGIYIRAGFIRRYVNVNQPSPLEHTEMLEQLRVLMYG
 EKHVAVAEVPGTGVDTPEDLERVAEMR"
 4106..4999
 /note="unnamed protein product; ORF_ID:0218#5"
 /codon_start=1
 /transl_table=11
 /protein_id="BAA35665.1"
 /db_xref="GI:4062490"
 /translation="MEQLRAELSHLLEKLSRIECVNEKADTALWALYDSQGNPMLP
 ARSFSTPGKAROLAKTMTLARSCTVMTPIYGVMTHERHPGPDVLLLRMRGVSVEA
 PARTPEREQLKDQIUEALLAHRQDSRCGVADNTQENFSPSWRQHVLEVLTILN
 QFNNTGLTMDQRIILFRTRIECPALFEGFNDNCVLHGNFCLSRMLKSDRSOQLLAW
 GFGMLWAPREYELFRLMDNSLAEDLLMSYLOQAPVAESFIWRRMLYVLWDEVAQLVN
 TGRFRRNFDLASKSILLPWL"
 complement (4976..5755)
 /gene="ycbC"
 complement (4976..5755)
 /gene="ycbC"
 /note="ORF ID:0218#6
 similar to SwissProt Accession Number P36565"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein in kdsB-mukF intergenic
 region."
 /protein_id="BAA35666.1"
 /db_xref="GI:4062491"
 /translation="MLFTLKVIYGNMLPLPLMLLIIGALLWFSRFQKTGKIFIS
 IGWALLLLSLQPVADRLRLPIESTYPTWNNQKVDYIVVLGGYTNWQWAPSSNLI
 NNSLPRLNEGIRLWRENGSKLIFTGGVAKNTVSTAEGARVAQSLGVPRQIITLD
 LPKDTSEEAAYKQAIQDAPFLVTSASHLPRAMIFFQOEGNLPLPAPANQLAIDSPL
 NPERAIKSPVWMLHSDRVGYETLGRIMQWLKSGSGEPQRE"
 5891..6676
 /gene="smtA"
 5891..6676
 /gene="smtA"
 /note="ORF ID:0218#7
 similar to SwissProt Accession Number P36566"
 /codon_start=1
 /transl_table=11
 /product="SmtA protein."
 /protein_id="BAA35667.1"
 /db_xref="GI:1651445"
 /translation="MQDRNFDDIAEKFSRNIYGTTKQQLRQAILWQDLRVLAEQGPQ
 KLRVLDAGGEGQTAIKWAERGHQVILCDLSAQMDIDRAKQAAEKGVSDNMFHCAA
 QDVASHLETFPVDLILFHAVLEWADPRSVLQTLMSVLRPGVLSLMFYNAHGLLHMNM
 VAGNDFVOAGMPKKKKTLSPDYPRDPAQVTLWLEEGAGWQIMGKTGVRVFDHLYLREK
 HQQRDCYALLELETRYCRQEPYITLGRYIHTVARTKPSQSKV"
 6673..7995
 /gene="kicB"
 /note="synonym: mukF"
 6673..7995
 /gene="kicB"
 /note="ORF_ID:0218#8

Alignment Scores:

Pred. No.: 2,09e-147 Length: 20604
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservativeness: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x D90730 (1-20604)

Qy	1	MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	18063	ATGTTTGAGAACATTACCGCCCTCTGCCGACCGCTGCGGCTGCGCCATCTGTT	18004
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTylyAspGluThr	40
Db	18003	CGTCCGATGACGTCGCCGCAAAATTAACCTCGGATTTGGTGTCTATAAGATGAGAG	17944
Qy	41	GlyLysThrProValLeuThrSerValIysLysAlaGluGlnTyLeuLeuGluAsnGlu	60
Db	17943	GGCAAAACCCGGTACTGACACAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAATGAA	17884
Qy	61	ThrThrLysAsnTyLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
Db	17883	ACCACCAAAATTAACCTCGGCATTGACGCGATCCCTGAAATTTGGTGGCTGCACCTCAGGAA	17824
Qy	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
Db	17823	CTGCTGTTTGGTAAGGTAGCGCCCTGATCANTGACAAACGTGCTCGCACGCGACAGACT	17764
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal	120
Db	17763	CGCGGGGCGACTAGCGGCGACTAGCGTGGCTGGCGATTTCTGGCAAAAATACACAGCGTT	17704
Qy	121	LysArgValTrrValSerAsnProSerTrrProAsnHisLysSerValPheAsnSerAla	140
Db	17703	AAGCGTGTGGGTGAGCAACCAACAGCTGGCGCAACCAATGAGAGCGTCTTTAACTCTGCA	17644
Qy	141	GlyLeuGluValArgGluTyAlaTyTrAspAlaGluAsnHisThrLeuAspPheAsp	160
Db	17643	GGTCTGGAAGTTCTGTAATACGCTTATATGATGGGAAATACACACTCTTGACTTCGAT	17584
Qy	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
Db	17583	GCACTGATTAACAGCGCTGAATGAAGCTCAGGCTGGCGAGTGTGCTGCTCCATGGCTGC	17524
Qy	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyTrGlnThrLeuAlaGlnLeu	200
Db	17523	TGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCACACTC	17464
Qy	201	SerValGluLysGlyTrrLeuProLeuPheAspPheAlaTyTrGlnGlyPheAlaArgGly	220
Db	17463	TCCGTTGAGAAAGCTGGTTACCGCTGTTGACTTCGCTTACCAGGGTTTGGCCGCTGT	17404
Qy	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaIleMetHisLysGlyLeuIleVal	240
Db	17403	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGGGCTATGCAATAAGAGCTGATTGTT	17344
Qy	241	AlaSerSerTySerLysAsnPheGlyLeuTyTrAsnGluArgValGlyAlaCysThrLeu	260
Db	17343	GCCAGTTCTTCTCTAAAAAATTTGGCCTGTACAAACGAGCGGTGTTGGCGCTGTACTCTG	17284
Qy	261	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
Db	17283	GTGCTGCGGACAGTGAACCGCTTATCGCGATTCAGCCAAATGAAGCGCGATTTCG	17224
Qy	281	AlaAsnTySerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300
Db	17223	GCTAACTACTCTAACCCACGACACACGCGCTCTCTGTTGTTGCCACCATCTGTAGCAAC	17164
Qy	301	AspAlaLeuArgAlaIleTrrGluGlnLeuLeuThrAspMetArgGlnArgIleGlnArg	320
Db	17163	GATCGCTTACGTGGATTTGGGAACAGAGCTGACTGTATATGCGCCAGCGTATTACGCGT	17104

Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGlnLysGlyAlaAsnArgAspPheSerPhe	340
Db	17103	ATGGCTCAGTTGTTTCGTCAATACGCTGAGGAAAAAGCGCAAAACCGGACTTCAGCTTT	17044
Qy	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	17043	ATCATCAAAACAGAACGCGCATGTTCTCTTCAGTGGCCTGACAAAAGAACAGTGTGCGT	16984
Qy	361	LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet	380
Db	16983	CTGCGCGAGAGTTTGGGCTATATGCGGTGCTTCTGCTCGGTAATATGTGGCCGCGATG	16924
Qy	381	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	16923	ACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCTG	16876

RESULT 11

U00096_09/c

WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (10 of 47) of U00096 from base 900001 (U00096 Escherichia coli K-12 MG1655

Alignment Scores:

Pred. No.: 1.53e-146 Length: 110000

Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x U00096_09 (1-110000)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
DB 84932 ATGTTTGAGAACATTACCGCCGCTCTGCGGACCCGATCTGGGCTGGCGATCTGTT 84873
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValIysAspGluThr 40
DB 84872 CGTCCGATGAACGTCGCGCAAAATTAACTCCGGATTGGTGTCTATAAGATGAGACG 84813
QY 41 GlyIleThrProValLeuThrSerValIysIleAlaGluGlnIleLeuGluAsnGlu 60
DB 84812 GGCAAAACCCCGGTACTACCAAGCGTGAAGAGCGTGAACAGTATCTGCTCGAAAATGAA 84753
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
DB 84752 ACCACAAAATTACTCGGCATTGACGGCATCCCTGGAATTTGGTGGCTGCACTCAGAA 84693
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
DB 84692 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCGTCGACGACAGACT 84633
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
DB 84632 CCGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGGCAAAAATACACGCGTT 84573
QY 121 LysArgValTropValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
DB 84572 AAGCGTGTGGTGAGCAACCAAGCTGGCGCAACCAAGAGCGTCTTTAACTCTGCA 84513
QY 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
DB 84512 GGTCTGGAAGTTCGTGATACGCTATTATGATGCGGAATATCACACTCTTGACTTCAT 84453
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
DB 84452 GCACGTATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTCTGTTCCATGGCTG 84393
QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTropGlnThrLeuAlaGlnLeu 200
DB 84392 TGCCATAACCAACCGGTATCGACCTACGCTGGAAACCAATGCAACACTGGCAACTC 84333
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
DB 84332 TCCGTTGAGAAGGCTGTTACCGCTGTTTGNCTTCGTTACCGGGTTTGGCCGTGTT 84273
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
DB 84272 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTCGGCGGTATGCATAAAGAGCTGATTGTT 84213
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
DB 84212 GCCAGTTCCTACTCTAAATAAATTTGGCCTGTACACAGCGGTGGCGCTGTACTCTG 84153
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
DB 84152 GTTGTGCGGACAGTGAACCGTTGATCGCGATTACGCCAAATGAAGCGGCGATTCCG 84093
QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaThrIleLeuSerAsn 300
DB 84092 GCTAACTACTCTAACCCACGACACAGCGCGCTCTCTGTTGTCACCATCTCTGAGCAAC 84033
QY 301 AspAlaLeuArgAlaIleTropGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
DB 84032 GATCGTTTACGTGGCATTTGGAAACAAGAGCTGACTGATATGCGCGACGATTCAGCGT 83973
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340

DB 83972 ATGCGTCAAGTTGTTGTCGAATACGTCGAGGAAAAGCGCAAAACCGGACTTCAGCTTT 83913
QY 341 IleIleLysGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArg 360
DB 83912 ATCATCAACAAAGAACGGCATGTTCTCTTCAGTGGCCCTGACAAAAGAACAGTGTGCGT 83853
QY 361 LeuArgGluGluPheGlyValIleAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
DB 83852 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTCTTCGTGCGTAAATGTCGCGGATG 83793
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
DB 83792 ACACCAATAACATGCTCCGCTGTCGGAAGCATTTGTGGCAGTCTG 83745

RESULT 12

AE005174_11/c

WPCOMMENT

Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name	Begin	End
AE005174_00	1	110000
AE005174_01	100001	210000
AE005174_02	200001	310000
AE005174_03	300001	410000
AE005174_04	400001	510000
AE005174_05	500001	610000
AE005174_06	600001	710000
AE005174_07	700001	810000
AE005174_08	800001	910000
AE005174_09	900001	1010000
AE005174_10	1000001	1110000
AE005174_11	1100001	1210000
AE005174_12	1200001	1310000
AE005174_13	1300001	1410000
AE005174_14	1400001	1510000
AE005174_15	1500001	1610000
AE005174_16	1600001	1710000
AE005174_17	1700001	1810000
AE005174_18	1800001	1910000
AE005174_19	1900001	2010000
AE005174_20	2000001	2110000
AE005174_21	2100001	2210000
AE005174_22	2200001	2310000
AE005174_23	2300001	2410000
AE005174_24	2400001	2510000
AE005174_25	2500001	2610000
AE005174_26	2600001	2710000
AE005174_27	2700001	2810000
AE005174_28	2800001	2910000
AE005174_29	2900001	3010000
AE005174_30	3000001	3110000
AE005174_31	3100001	3210000
AE005174_32	3200001	3310000
AE005174_33	3300001	3410000
AE005174_34	3400001	3510000
AE005174_35	3500001	3610000
AE005174_36	3600001	3710000
AE005174_37	3700001	3810000
AE005174_38	3800001	3910000
AE005174_39	3900001	4010000
AE005174_40	4000001	4110000
AE005174_41	4100001	4210000
AE005174_42	4200001	4310000
AE005174_43	4300001	4410000
AE005174_44	4400001	4510000
AE005174_45	4500001	4610000
AE005174_46	4600001	4710000
AE005174_47	4700001	4810000
AE005174_48	4800001	4910000
AE005174_49	4900001	5010000
AE005174_50	5000001	5110000
AE005174_51	5100001	5210000
AE005174_52	5200001	5310000

Db	103563	GCTAACTACTCTTAACCCACCAGCACACGGCGCTTCTGTGTGGCCACCATCTGAGCAAC	103504
Qy	301	AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
Db	103503	GATGCGTTACGTGGGATTTGGGAACAAGAGCTGACTGATATGCCCCAGCGTATTTCAGCGT	103444
Qy	321	MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe	340
Db	103443	ATGCGTCAGTTGTCGTCAATACGTCAGGAAAAAGCGCAACCCGCGACTTCAGCTTT	103384
Qy	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
Db	103383	ATCATCAAAACAGAACGCGCATGTTCTCCCTTCAGTGGCTGACAAAGAAACAAGTGTGCGT	103324
Qy	361	LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet	380
Db	103323	CTGGCGAAGAGTTTGGCGGTATGCTTGTCTTGGTCGGTAACGTGGCGGGGATG	103264
Qy	381	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	103263	ACACCAGATAACATGGCTCCGCTGTGGAAGCGATTGTGGCAGTGTG	103216
RESULT 13			
AE005174_12/c			
WPCOMMENT			
Sequence split into 56 fragments LOCUS AE005174 Accession AE005174			
Fragment Name Begin End			
AE005174_00	1	110000	
AE005174_01	100001	210000	
AE005174_02	200001	310000	
AE005174_03	300001	410000	
AE005174_04	400001	510000	
AE005174_05	500001	610000	
AE005174_06	600001	710000	
AE005174_07	700001	810000	
AE005174_08	800001	910000	
AE005174_09	900001	1010000	
AE005174_10	1000001	1110000	
AE005174_11	1100001	1210000	
AE005174_12	1200001	1310000	
AE005174_13	1300001	1410000	
AE005174_14	1400001	1510000	
AE005174_15	1500001	1610000	
AE005174_16	1600001	1710000	
AE005174_17	1700001	1810000	
AE005174_18	1800001	1910000	
AE005174_19	1900001	2010000	
AE005174_20	2000001	2110000	
AE005174_21	2100001	2210000	
AE005174_22	2200001	2310000	
AE005174_23	2300001	2410000	
AE005174_24	2400001	2510000	
AE005174_25	2500001	2610000	
AE005174_26	2600001	2710000	
AE005174_27	2700001	2810000	
AE005174_28	2800001	2910000	
AE005174_29	2900001	3010000	
AE005174_30	3000001	3110000	
AE005174_31	3100001	3210000	
AE005174_32	3200001	3310000	
AE005174_33	3300001	3410000	
AE005174_34	3400001	3510000	
AE005174_35	3500001	3610000	
AE005174_36	3600001	3710000	
AE005174_37	3700001	3810000	
AE005174_38	3800001	3910000	
AE005174_39	3900001	4010000	
AE005174_40	4000001	4110000	
AE005174_41	4100001	4210000	
AE005174_42	4200001	4310000	
AE005174_43	4300001	4410000	
AE005174_44	4400001	4510000	
AE005174_45	4500001	4610000	


```

/gene="dacC"
/locus_tag="c0924"
/EC_number="3.4.16.4"
/function="enzyme; Murein sacculus, peptidoglycan"
/note="D-alanyl-D-alanine carboxypeptidase fraction C;
Escherichia coli K-12 ortholog: b0839; Escherichia coli
O157:H7 ortholog: z1066"
/codon_start=1
/transl_table=11
/product="Penicillin-binding protein 6 precursor"
/protein_id="AAN79397.1"
/db_xref="GI:26107214"
/translation="MDTRVAFMTQYSSILRLGAAGSFLFLFAPTAFAAEQIVAPSV
DAWATWLDYAGKVLAEAGNDEKLDPAKSLTKIMTSYVVGQALKADKIKLTDMTVVGK
DAWATGSSVWFLKPGQDVADLNKGVITQSGNDACIALADYVAGSQESFIF
LMNGVAKGLTNTTFTVTHGLDAPQGFSTARDMALLGKALIDHVPPEYVATIKKEFT
FNKIQPNRLLWSSNLNVDMKMTGTAGAGYNLVASATQDMRLISVVLGAKTDRI
RNESEKLTWGFREFETPIPKDATFTVQRWFGDKSEVNLGAGEAGSVTIPKQL
KNLKASYLTLEPQLAPLKGGVGTIDFQLNGKSIQRPLIVMENVEGGFFGRMWD
FVMKFFHQWFGSWS"
/complement(1411..2169)
/gene="deoR"
/locus_tag="c0925"
/complement(1411..2169)
/gene="deoR"
/locus_tag="c0925"
/function="regulator; Salvage of nucleosides and
nucleotides"
/note="Escherichia coli K-12 ortholog: b0840; Escherichia
coli O157:H7 ortholog: z1067"
/codon_start=1
/transl_table=11
/product="Deoxyribose operon repressor"
/protein_id="AAN79398.1"
/db_xref="GI:26107215"
/translation="METREERIRGQLQELKRSLKHDKDAALGVSEMTIRDLNN
HSAPVLLGGYIVLPRASHYLLSDQSRVVEERRAAKLAATLVEPDQLFFDCGT
TPVLTIEADNEIPPTAVCYSLNTFLAKEPKHCFACGGEFHASNAIFKPIDFOQT
LNNFCPDIAFYSAAGVHYSKGCATCNLELPVKHWMSMAQKHLVVDHSFEGKVRPA
RMGLDKRFDIVSDCCPEDEYVYKATQRIKLMY"
/complement(2227..2823)
/gene="ybjG"
/locus_tag="c0926"
/complement(2227..2823)
/gene="ybjG"
/locus_tag="c0926"
/note="Escherichia coli K-12 ortholog: b0841; Escherichia
coli O157:H7 ortholog: z1068"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjG"
/protein_id="AAN79399.1"
/db_xref="GI:26107216"
/translation="MLENLNLSFLINATPDSAPWMSLAIFIAKDLITVVPILAAV
LWLWGLVITFEALAFICWRLWNGSLMLVLAVLVIASVYLVGHVWDLMLGGLLAGMIG
CLSAQIIWQAMGHKLYQRLQSWRFPCALPIRKWVRD"
/locus_tag="cmr"
/locus_tag="c0927"
/gene="cmr"
/locus_tag="c0927"
/function="transport; Transport of small molecules: Other"
/note="Chloramphenicol resistance pump cmr; Escherichia
coli K-12 ortholog: b0842; Escherichia coli O157:H7
ortholog: z1069"
/codon_start=1
/transl_table=11
/product="Multidrug translocase mdfA"
/protein_id="AAN79400.1"
/db_xref="GI:26107217"
/translation="WRDYWRRCNQNKLASGRQALLPPLCLVLYESTYIGND

```

```

MIQCMILAVVEQYQAGIDWVTSMTAYLAGCMFLQWLLGCLPLSDRIGRPVMLAGVVFV
IVTCLAILLAQNIQFLLRFLQGISLCFICAVGAYAAIQESFEEAVATIKITALMANVA
LIAPLLGLVAQNIHVLWPEGMFVFAALAAISFFGLQRPMTATVATIGEKLSKEL
GRDYLKLVKNGRFVAGALAGFVSLPIAWIAQSPITIIITGEQLSSYSGVGLQVPIFG
ALAGNLLARLTSSRTVRSLLIIMGWPIMIGLLVAAATVTSSSHAYLWMTAGLSIYA
FGIAGNAGLVRLTIFASDSKSTVSAMGMQLMLIFTVVGIEISIKHAWLNGGGLPNL
FNLVNGILWLSLWIFLKDQMGNSHEG"
/complement(4381..4665)
/gene="ybjH"
/locus_tag="c0928"
/complement(4381..4665)
/gene="ybjH"
/locus_tag="c0928"
/note="Escherichia coli K-12 ortholog: b0843; Escherichia
coli O157:H7 ortholog: z1070"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjH precursor"
/protein_id="AAN79401.1"
/db_xref="GI:26107218"
/translation="WIMKNCILLGALLMGFTGVAMAQSVTVVPSGYKVVVVPDSVSV
POAVSVATPQTVVAAPAPAPAYRPHVPVRHLASVGEQWVIEHQIDDDHH"
/complement(4751..5569)
/locus_tag="c0929"
/complement(4751..5569)
/locus_tag="c0929"
/note="Escherichia coli K-12 ortholog: b0844; Escherichia
coli O157:H7 ortholog: z1071"
/codon_start=1
/transl_table=11
/product="Protein ybjI"
/protein_id="AAN79402.1"
/db_xref="GI:26107219"
/translation="WMSIKLIAVDMDGTFLSDQYNERERMAOYQOMKAQGRFVVA
SGNQYQLISFFPEIANEIAFVAENGWVSEKDVFNGLSKDAPATVVEHLLTRPE
VEIIACGKNSAYTLKKYDDAMTKVAEMTYHRLVYVDNFNLEIDFFKGLNLSDELIP
QVQALHEAIGDIMVPHVTGNGSIDLIIIPGVHKANGRLQOLKMGIDSEVVVFDGGG
NDIEMLRQAGSFAMENAGSAVAAAKYRAGSNRNREGVLDVDRKHEAPFKQ"
/complement(5566..6837)
/gene="ybjJ"
/locus_tag="c0930"
/complement(5566..6837)
/gene="ybjJ"
/locus_tag="c0930"
/function="putative regulator"
/note="Escherichia coli K-12 ortholog: b0845; Escherichia
coli O157:H7 ortholog: z1072"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjJ"
/protein_id="AAN79403.1"
/db_xref="GI:26107220"
/translation="MCTNVHSLANVFPPLLITWMTVNSRNALKRRTWALFMPFFFL
PGLLMASWATTPAIRDILSVIAEMGVLFGLSIGSMGILCSAWLVKRGTRNVIL
VTMSGALIMILSLALWTSPLFVGLVFGASFGSAEVAIVNVEGAVERENKTV
LPMHGFYSLCTLAGVGMALTAFGVPAVTVHLLAALVGIAPVIAIQAIPIDGTGKN
NADCTQGEKGVPEYRDIQLLIGVVLAMAFSGSANDMLPMLWDGHEGFSPTSSSL
IYAGTFLGTVGRFIMGFDIHSRVAVVRASALMGALIGLIIIFVDSANWAGVSVVL
WGLGASIGPLTISAADTPDAPTRVSVVATGYLAFVLGPPLLIGLGEHYGLRSAM
LVVLALVILAAIVAKAVAKPDTKTQTAMENS"
/locus_tag="c0931"
/locus_tag="c0931"
/gene="ybjK"
/locus_tag="c0931"
/function="putative regulator"
/note="Escherichia coli K-12 ortholog: b0846; Escherichia
coli O157:H7 ortholog: z1073"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ybjK"
/protein_id="AAN79404.1"

```

```

/db_xref="GI:26107221"
/translation="MYTIAERGNMRRANDPORREKLIQATLEAVKLYGIHAVTHRKI
AALAGVPVGSMTYFSGIDELLESFMSRQYQAFPSVDSDPAGCAQAITDMI
YSSQVATPDNNELMYQYALASRKPLIKTVQNMQRSSQOTLEQWFFPGTARALDAFI
EGMTLHFVTDKRLPSRBEILRMVERIAG"
complement(7873..8904)
/gene="intn"
/locus_tag="c0932"
complement(7873..8904)
/gene="intn"
/locus_tag="c0932"
/function="enzyme; Integration, recombination Phage or
Prophage Related"
/notes="Escherichia coli O157:H7 ortholog: z2966"
/codon_start=1
/transl_table=11
/product="Integrase for prophage"
/protein_id="AAN79405.1"

```

Alignment Scores:

```

Pred. No.: 4,27e-145 Length: 301276
Score: 2033.00 Matches: 392
Percent Similarity: 100.0% Conservative: 4
Best Local Similarity: 99.0% Mismatches: 0
Query Match: 99.4% Indels: 0
DB: 1 Gaps: 0

```

US-10-673-786A-2 (1-396) x AE016758 (1-301276)

```

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 135038 ATGTTTGAGAACATTACCGCGCTCTCGCCACCGGATTCTGGCGCTGGCGGATCTGTTT 134979
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 134978 CGTCCGATGACGTCGCGGCAAAATTAACTCTCGGATTGGTGCTCTATAAGATGAGACG 134919
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 134918 GGTAACCCCGGTACTGACCAGCGTGAAAAGGCTGAACAGTATCTGCTCGAAAATGAA 134859
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 134858 ACCACCAAAAATTACCTCGGATTTGACGGCTCTCCCTGAAATTTGGTCGCTGACCTCAGGAA 134799
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 134798 CTGCTGTTTGGTAAGGTAGGCGCTGATCAATGACAAACGCTGTCGACCGGACAGACT 134739
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 134738 CCGGTGGCACTGGCGCACTACGGGTGGCTGCCGATTTCTGGCAAAAATAACACAGCGTT 134679
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 134678 AAGCGTGTGGGTGAGCAACCAAGCTGGCGCAACCACTAATAGAGCGCTTTAACTCTGCA 134619
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 134618 GGTCTGGAAGTTCGTGAATACGCTATTATGATCGGAAATACACACTCTTGACTTTGAT 134559
Qy 161 AlaLeuIleAenSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 134558 GCACGTATTACAGCCTGAACGAGCTCAGGCTGGCGAGGTAGTGTGTTCCATGGGTGC 134499
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 134498 TGCCATAACCAACCGGATCGACCTACGCTGGAAACAAATGGCAGACACTGGCACAACTC 134439
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 134438 TCCGTTTGAAAGGCTGGTTTACCGCTGTTTACCTTCGCTTACCAAGGGTTTGGCCCGTGGT 134379

```

```

Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 134378 CTGGAAGAAGATGCTGAAGACTTGGCGCTTTCCGGCTCTGCAATAAGGAGCTGATTGTT 134319
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 134318 GCCAGTTTCCTACTCTAAAACCTTTGGCCTGTACAAACGAGCGTGTGGCGCTGTGACTCTG 134259
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 134258 GTTGCTGCTGACAGTGAACCTGTTGAACGCGCATTCAGCCAAATGAAAGCGCGGATTTCGT 134199
Qy 281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 134198 GCTAACTACTCTAACCCACCAGCACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 134139
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 134138 GATGCGTTACGTGGATTTCGGAAACAAGNACTGACTGATATGCCACGCTATTTCAGCGT 134079
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAenArgAspPheSerPhe 340
Db 134078 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAAGCGCAAAATCGCGACTTCAGCTTT 134019
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 134018 ATCATCAACAGAACCGCATGTTCTCCTTCAGTGGCTTGACAAAAGAACCAAGTCTCGCT 133959
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 133958 CTGGCGCAAGAGTTTGGCGTATATGCTGTTGCTTCTGTCGCTGAAACGTAGCTGGGATG 133899
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 133898 ACACCAAGATAACATGGCGCGCTGTGCAAGCGATTGTGGCAGTGTCTG 133851

```

Search completed: March 15, 2006, 16:49:38
Job time : 5484 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 15, 2006, 02:39:56 ; Search time 447 Seconds
(without alignments)
4736.188 Million cell updates/sec

Title: US-10-673-786A-1
Perfect score: 1191
Sequence: 1 atgtttgagaacattaccgc.....cgattggcagctgtaa 1191

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCFUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1191	100.0	1331	3	US-08-985-908-23
2	839	70.4	1209	3	US-09-489-039A-4902
3	458.2	38.5	1830121	3	US-09-557-884-1
4	458.2	38.5	1830121	3	US-09-643-990A-1
5	458.2	38.5	1830121	3	US-10-158-865-1
6	387.6	32.5	579	3	US-09-489-039A-4954
7	348.4	29.3	702	3	US-09-543-681A-3805
8	316.4	26.6	1509	3	US-09-252-991A-12968
9	316.4	26.6	1632	3	US-09-252-991A-12507
10	306.6	25.7	24417	2	US-08-846-762-1
11	305	25.6	1242	3	US-09-252-991A-12428
12	305	25.6	1329	3	US-09-252-991A-12162
13	289	24.3	1338	3	US-09-489-039A-1301
14	278.2	23.4	1185	3	US-09-489-039A-1274
15	276.4	23.2	1209	3	US-09-540-236-315
16	274.8	23.1	100848	3	US-09-596-002-39
17	263.6	22.1	1284	3	US-09-328-352-769
18	242.6	20.4	1260	3	US-09-543-681A-3123
19	218.8	18.4	2440	3	US-09-976-594-429
20	201.2	16.9	426	3	US-09-543-681A-3717
21	177.8	14.9	1941	3	US-09-220-132-65
22	177.8	14.9	1941	3	US-09-814-915A-35
23	177.8	14.9	1942	3	US-09-949-016-1216
24	176.2	14.8	2056	3	US-09-566-921-56

Sequence 59, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 4873, Ap
Sequence 4412, Ap
Sequence 12380, A
Sequence 145, App
Sequence 13293, A
Sequence 3374, Ap
Sequence 4436, Ap
Sequence 3375, Ap
Sequence 12333, A
Sequence 1235, Ap
Sequence 5, Appli
Sequence 658, App
Sequence 4632, Ap
Sequence 318, App
Sequence 952, App
Sequence 3376, Ap
Sequence 1603, Ap
Sequence 434, App

ALIGNMENTS

RESULT 1
US-08-985-908-23
; Sequence 23, Application US/08985908
; Patent No. 6004773
; GENERAL INFORMATION:
; APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI NASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,908
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-325659
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1331 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: JM109
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..1197
US-08-985-908-23

```
Query Match      100.0%; Score 1191; DB 3; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAAATTAACCGCGCTCTGCGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 60
DB 10 ATGTTTGAGAAATTAACCGCGCTCTGCGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 69
QY 61 CGTGCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGAG 120
DB 70 CGTGCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGAG 129
QY 121 GCAAAACCCGCTACTGACACGGTGAAAGAGCTGACAGTATCTGTCGAAATGAA 180
DB 130 GCAAAACCCGCTACTGACACGGTGAAAGAGCTGAAACAGTATCTGTCGAAATGAA 189
QY 181 ACCACCAAAATTAACCTCGGCATTTGACGGCATCCCTGAAATTTGGTGGCTGCACTCAGGAA 240
DB 190 ACCACCAAAATTAACCTCGGCATTTGACGGCATCCCTGAAATTTGGTGGCTGCACTCAGGAA 249
QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGACT 300
DB 250 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGACT 309
QY 301 CCGGGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTGCGCAAAATATACAGGTT 360
DB 310 CCGGGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTGCGCAAAATATACAGGTT 369
QY 361 AAGCGTGTGGTGAGCAACCAAGCTGGCCGCAACCAATGAGCGGCTTTAACTCTGCA 420
DB 370 AAGCGTGTGGTGAGCAACCAAGCTGGCCGCAACCAATGAGCGGCTTTAACTCTGCA 429
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTTTGACTTCGAT 480
DB 430 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTTTGACTTCGAT 489
QY 481 GCACTGATTAACAGCCTGAATGAGCTCAGGCTGGCGACGATGCTGTTTCCATGGCTGC 540
DB 490 GCACTGATTAACAGCCTGAATGAGCTCAGGCTGGCGACGATGCTGTTTCCATGGCTGC 549
QY 541 TGCCATAACCAACCGGTATCGACCTTACGCTCGAACCAATGGCAACACACTGGCACAACTC 600
DB 550 TGCCATAACCAACCGGTATCGACCTTACGCTCGAACCAATGGCAACACACTGGCACAACTC 609
QY 601 TCCGTTGAGAAAGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTTCGCGGTG 660
DB 610 TCCGTTGAGAAAGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTTCGCGGTG 669
QY 661 CTGGAAGAAGTCTGAAAGACTGCGCGCTTTTCGCGCTATGCAATGAAGCTGATTTT 720
DB 670 CTGGAAGAAGTCTGAAAGACTGCGCGCTTTTCGCGCTATGCAATGAAGAGCTGATTTT 729
QY 721 GCCAGTTCTTACTCTAAATACTTTGGCTGTGTAACAGAGCGTGTGGCGCTTGTACTCTG 780
DB 730 GCCAGTTCTTACTCTAAATACTTTGGCTGTGTAACAGAGCGTGTGGCGCTTGTACTCTG 789
QY 781 GTTGCTGCCGACAGTGAACCGTTGATPCGCGCATTCAGCCAAATGAAGCGGCGATTCGC 840
DB 790 GTTGCTGCCGACAGTGAACCGTTGATPCGCGCATTCAGCCAAATGAAGCGGCGATTCGC 849
QY 841 GCTAACTACTCTAACCCACAGACACGGGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
DB 850 GCTAACTACTCTAACCCACAGACACGGGCGCTTCTGTTGTCACCATCTCTGAGCAAC 909
QY 901 GATGCGTTAGTGGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTTACGCGT 960
DB 910 GATGCGTTAGTGGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTTACGCGT 969
QY 961 ATGCGTCAGTTGTTGTCGTAATAGCTGCAGGAAAAAGCGCAAAACCGCGACTTCAGCTTT 1020
DB 970 ATGCGTCAGTTGTTGTCGTAATAGCTGCAGGAAAAAGCGCAAAACCGCGACTTCAGCTTT 1029
QY 1021 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT 1080
```

```
DB 1030 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTGTGCGT 1089
QY 1081 CTGCGCGAAGAGATTTGGCGTATATGCGGTGCTTCTGTCGCGTAATGTCGCCGGATG 1140
DB 1090 CTGCGCGAAGAGATTTGGCGTATATGCGGTGCTTCTGTCGCGTAATGTCGCCGGATG 1149
QY 1141 ACACAGATAAATGCTCCGCTGCGAAGCATTTGTGCGAGTGTCTGTAA 1191
DB 1150 ACACAGATAAATGCTCCGCTGCGAAGCATTTGTGCGAGTGTCTGTAA 1200

RESULT 2
US-09-489-039A-4902
; Sequence 4902, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4902
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4902

Query Match      70.4%; Score 839; DB 3; Length 1209;
Best Local Similarity 81.5%; Pred. No. 6.7e-273;
Matches 971; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAAATTAACCGCGCTCTGCGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 60
DB 19 ATGTTTGAGAAATTAACCGCGCTCTGCGCGACCCGATTCCTGGCCCTGGCCGATCTGTTT 78
QY 61 CGTGCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGAG 120
DB 79 CGTGCCGATGAACGTCCTGGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGAG 138
QY 121 GCAAAACCCGCTACTGACAGCGTGAAGAGCTGAAACAGTATCTGTCGCAAAATGAA 180
DB 139 GGTAAACCGCGCTTCTGACAGCGTCAAAAAAGCAGACGATCTGCTGGAATAATGAA 198
QY 181 ACCACCAAAATTAACCTCGGCATTTGACGCGATCCCTGAAATTTGGTGGCTGCACTCAGGAA 240
DB 199 AGACTTAAACCTATCTGGGCATCGATGGTATTTCTGAAATTTGGTGGCTGCACTCAGGAG 258
QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGCAAGGACAGACT 300
DB 259 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGCAAGGACAGACT 318
QY 301 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGGATTTCTGCGCAAAATATACAGGTT 360
DB 319 CCGGGGGGCACTGGCGCTGCTGCTGCGCGAGACTTCTCGCCAAAAACACCGAGCTG 378
QY 361 AAGCGTGTGGTGAGCAACCCAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 420
DB 379 AAACGTTGTGGGTAAAGTATCCGAGCTGGCCGCAACCAATGAAGCGTATTTACCTCTGCC 438
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAAAATCACACTCTTGACTTCGAT 480
DB 439 GGGCTGGAAGTTCGCGCAATACGATACGACGCGGCTAAACACGCGCTGGACTTTGAT 498
QY 481 GCACTGATTAACAGCCTGAATGAGCTCAGGCTGGCGACGATGCTGTTTCCATGGCTGC 540
DB 499 GGTCTGCTGCCAGCGCTGAAACGAAAGCCAGCGGCGGACGTTGGTACTGTTTCCACGGCTGC 558
QY 541 TGCCATAACCAACCGGTATCGACCTTCAGCCCTACGCTGGAACCAATGGCAACAACTGGCAACTC 600
```

Db 559 TGCCACACCCGACCGGTATCGATCCGACGCTCGATCAGTGGCAGCAGCTGGCGAGCTG 618
Qy 601 TCCGTTGAGAAAGGCTGTTACCGCTGTTTACCTTGGCTTACCTAGGGTGGTGGCGGTGGT 660
Db 619 TCGGTGGAAGGCTGCTACCGCTGTTTCGATTTCCGCTACCAAGGGCTTCGCCGCGGT 678
Qy 661 CTGGAAGAAGATGCTGAAGAGCTGCGCGCTTTTCGCGGCTATGCAATAAGAGCTGATTGTT 720
Db 679 CTGGAAGAAGATGCTGAAGGCTGACGCGCTTTTGGCTCTACATAAAGAGCTGCTGTC 738
Qy 721 GCCAGTTCTACTCTAAAGGCTTGGCGCTGTACACGAGCGTGGTGGCGCTTGTACTCTG 780
Db 739 GCCAGTTCTACTCTAAAGGCTTGGCGCTGTACACGAGCGCTGGCGCGCTGCACTCTG 798
Qy 781 GTTGTGCGCAGTGAAGCGTGTGATCGCGCATTTCAAGCAAAATGAAGGCGCGATTCGC 840
Db 799 GTCCGCGGATCAGGAGACTGTAGACCGCGCTTCAGTCAGATGAAGTCGGTGATCCGC 858
Qy 841 GCTAACTACTTAACCCACAGCAGCAGCGCGCTTCTGTTGTCACCATCTCGAGCAAC 900
Db 859 GCCAACTACTGAAACCGCTCGCATGCGCGCTCCGTTGGTGGCCACCATTTCTGAGCAAC 918
Qy 901 GATCGGTTACGTGGGATTTGGGAACAAGAGCTGATGATATGCGCGCAGCGTATTCAGCGT 960
Db 919 GATCGGTTACGTGGGATTTGGGAGCAGGAATGACCGATATGCGCGCAGCGCATCCAGCGT 978
Qy 961 ATGCGTCAGTTGTCGTCATACCTGTCAGGAGAAAGCGCAAAACCGCGACTTCAGCTTT 1020
Db 979 ATGCGTCGTTGTCGTCATACCTGTCAGGAGAAAGCGCGCGCGACTTCAGCTTT 1038
Qy 1021 ATCATCAACAGAGCGGATGTTCTCTTTCAGTGGCGCTGACAAAGAACAAAGTCTGCGT 1080
Db 1039 ATCAGCAGCAGAGCGGATGTTCTCTTTCAGCGCGCTGACTAAAGCAGGCTGCTGGC 1098
Qy 1081 CTGCGGAGAGTTTGGGATATGCGGTTGCTTCTGTCGCGTAAATGTCGCGGAGT 1140
Db 1099 CTGCGTGAAGATTTCGCGCATCTATGCGGTAGCTTCCGAGCATCAACGTCGCGCGGATG 1158
Qy 1141 ACACGATACATGCTGCGCTGTCGAGCGGATTTGGCAGTCTGTAA 1191
Db 1159 ACGCTTGACATATGGCGCGCTGTGGAAGCCATCGTCGCGCTACTGTAA 1209

RESULT 3

US-09-557-884-1

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of

; the Haemophilus influenzae Rd Genome, Fragments

; Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/557,884

; FILING DATE: 25-Apr-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/476,102

; FILING DATE: JUN-5-1995

; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 38.5%; Score 458.2; DB 3; Length 1830121;
Best Local Similarity 61.5%; Pred. No. 5.5e-142;
Matches 733; Conservative 0; Mismatches 458; Indels 0; Gaps 0;
Qy 1 ATGTTTTCAGAAACATTACCGCGCTCTCTGCGACCCCGATTCTGGGCGCTGGCGGATCTGTTT 60
Db 1684147 ATGTTTTCAGAAACATTCAAGCGGCGACCGATCCATCTTAGGCTTAGCGGAAGCAATTT 1684205
Qy 61 CGTCCCGATGAACGTCCCGGCAAAATTAACCTCGGATTCGGTGTCTATAAAGATGAGACG 120
Db 1684207 AAATCCGAAATTCGCGAAATATAATCAATTTGGGTATTTGGCGTTTATAAAGATGCGCAA 1684266
Qy 121 GGCMAAACCCCGGTACTGACGAGGTGAAAAGGTGAACAGTATCTGCTCGAAATGAA 180
Db 1684267 GGCACAAACCCCAATTATGACGCGGTAAAAAGAGCCGAAACGATATTATTGATAAGGAA 1684326
Qy 181 ACCACCAAAATTTACCTTCGGCATTTGACGCGATCCCTCAATTTGGTCTGCTGCACTCAGGAA 240
Db 1684327 AAAACCAAGAAATTTCTGACTATCGATGTTTGGGATTTATACGAAACAAACAAAGCA 1684385
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCAGCGCAGACT 300
Db 1684387 CTCCTTTTCGTTAAAGATTCTGAAGTCTCAATCTAATCGAGCAAGAACAGTACAAAGT 1684446
Qy 301 CCGGGGGGCACTGGCGCACTACGGTGGCTGCGGATTTCTTCGCAAAAAATACCAGCGTT 360
Db 1684447 TTAGCGCGAAACAGGTGCATTTACGCATTTGCGCGAGAAATTTTAAACGCCAAACTAAAGCA 1684506
Qy 361 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCATAGAGCGCTTTAACTCTGCA 420
Db 1684507 CAAATGTTTGGWTGAGCAACCACTTTGCCAAACCAATGCGATTTTCAATGCTGTC 1684566
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATTTATGATCGCGGAAATCACACTCTTGACTTCGAT 480
Db 1684567 GGTATGACCATTCGTGAATATCGTTATTTATGATGCTGAACGCAAGGCCCTTTGATTGGGAA 1684626
Qy 481 GCATGATTAACAGCCTGAATGAAGTTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 540
Db 1684627 CACTTATTAGAAGATTTAAGCCAAAGCAAGCGGATGTTGTTGCTTTTACACGGTTGT 1684686
Qy 541 TGCCTAATCCCAACCGGTATCGACCCCTACCTCGAACAATGGCAAAACACTGGCAACAATC 600
Db 1684687 TGCCTAATCCGACTGGTATTGACCCCTACTCCAGAAACAATGGCAAGAAATATAGCCGCACTT 1684746
Qy 601 TCCGTTTGAGAAAGCGTGTGTACCGCTGTTTGGACTTTGCTTTACAGGCGGATTTTGGCCGCTGGT 660
Db 1684747 TCAGCTAAATAATGGTGGTTGGCTCTTTGACCTTTTGGCTTTATCAAGGTTTAGCCAAACGGA 1684806
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCAATAAAGAGCTGATTGTT 720
Db 1684807 TTAGATGAAGATGCTTATGGTTTACGTCGCTTTTTCAGCAAAACCAACAAAGAAATATTAGTG 1684866
Qy 721 GCCAGTTCTACTCTTAAAGAACTTTGGCGCTGTACACGAGCGCTGTGCGCTTGTACTCTG 780
Db 1684867 GCGAGTTCTACTCTGAAAAAACTTTGGTTTATATAATGAACGTTGTTGTCATTTACCCCTT 1684926
Qy 781 GTTGTGCGCAGTGAAGCGGTTGATCGCGCATTTAGCGGCATTCAGCCAAATGAAGGCGGATTCGC 840


```

Qy 961 ATGGCTCAGTTGTTCTGTCATACGCTGCAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
Db 1685107 ATGGCTCAGTTGTTCTGTTCAATTAAGAAATATGTCGAGCAAGATTTCAGCTTT 1685166
Qy 1021 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCCCTGACAAAGAACAAAGTCTGCGT 1080
Db 1685167 ATCATTAACAAACGCGCATGTTCTCTTCAGTGGATTAACAGGGGAACAAAGTGGATCGT 1685226
Qy 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTTCTTCTGTCGCGTAATGTGCGCGGATG 1140
Db 1685227 TTAATAATGAATTTGCCATTACGCTGTTCTGTCGTCATCAACGTAAGTACGTAATC 1685286
Qy 1141 ACACGAGATACAGTCCGCTGTCGGAAGCGATTGTGGCAGTGCCTGTA 1191
Db 1685287 ACAGAAGATATATTCGCTATCTATGTGAAGATATCGTGAAGTACTTTAA 1685337

RESULT 5
US-10-158-865-1
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; FILE REFERENCE: P186P2CD1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g

```

;	NAME/KEY: misc_feature	
;	Query Match	38.5%; Score 458.2; DB 3; Length 1803121;
;	Best Local Similarity	61.5%; Pred. No. 5.5e-142;
;	Matches 733; Conservative	0; Mismatches 458; Indels 0; Gaps 0;
QY	1	ATGTTTGAGAACATTACCCCGCTCTCTGCCGACCCCGATTCTGGGCTGGCCGATCTGTTT 60
DB	1684147	ATGTTTGACATATCAAGCGGCACCGATCCAACTTAGGCTTAGCGAAGCATTT 1684206
QY	61	CTGCGCGATGAACGTCCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAGATGAGACG 120
DB	1684207	AAATCCGAACTCGGAAATAAATAATCAATTTGGGTATTTGGCGTTTATAAAGATGCGCAA 1684266
QY	121	GGCAAAACCCCGTACTGACCGCGTGAAGAGCTGAACAGTATCTGCTCGAATAATGAA 180
DB	1684267	GGCACAACCCCAATTATGACCGGGTAAAGAACGCAAAACGATTTTGTATAGGAA 1684326
QY	181	ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAATTTTGGTGCCTCGACTCAGGAA 240
DB	1684327	AAACCAAGAATTAICTGACTATCGATGTTTGGGATTTATAACGAACAAACAAAGCA 1684386
QY	241	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAGCTCTCGACGGCACAGACT 300
DB	1684387	CTCCCTTTTCGGTAAAGATTCTGAAGTCAATCAATCTAATCGAGCAAGAACAGTACAAAGT 1684446
QY	301	CGGGGGGCACCTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACACAGGTT 360
DB	1684447	TTAGCGGAACAGGTGCATTACGCAATTGGCGGAGAAATTTATTAAGCCCAACTAAGCA 1684506
QY	361	AAGCGTGTGTGGTGAGCAACCCCAAGCTGGCGGCAACCAATAAGAGCGCTCTTTAACTCTGCA 420
DB	1684507	CAAAATGTTTGGTTCAGCACACCAACTTGGCCAAACCAATGCGATTTTCAATGCTGTC 1684566
QY	421	GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
DB	1684567	GGTATGACCAATTCGTGAATATCGTTTATATGATGCTGAACGCAAGCCCTTGATGGGAA 1684626
QY	481	GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 540
DB	1684627	CACTTATTAGAAATTTAAGCCCAAGCAAGCAAGGCGATGTGGTGTCTTTTACAGGTTGT 1684686
QY	541	TGCCATAACCAACCGGTATCGACCCCTAGCTGTGAAACAATGGCAAACTTGGCACAACTC 600
DB	1684687	TGCCATAATCCGACTCGTATTCACCCCTACTCCAGAACAAATGGCAAGATTTAGCCGCACTT 1684746
QY	601	TGCGTTGAGAAAGCTGGTTACGCTGTTTGACTTCGCTTACCAGGTTTTCGCCGTGT 660
DB	1684747	TCAGCTAAAAATGGTTGGTTGCCACTCTTTGACTTTTGTCTTATCAAGGTTTAGCCAAACGGA 1684806
QY	661	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGCGCTATGATTAAGAGCTGATTGTT 720
DB	1684807	TTAGATGAAGATGCTTATGTTTACGTGCTTTTTCGAGCAAAACCAAGAAATTTATAGTG 1684866
QY	721	GCCAGTTCCTACTCTAAAAAATTTGGCCTGTACAACGAGCGTGTGGCGTGTGTACTCTG 780
DB	1684867	GCGAGTTCATTCTCGAAAACTTTGGTTATATAATGAACGCTGTGGTGCAATTTACCCTT 1684926
QY	781	GTTGCTGCCGACAGTCAAAACCGTTGATTCGCGCATTCAGCCAAATGAAGCGGCGATTCCG 840
DB	1684927	GTGGCAGAAAAATGCAGAAATTCATCAACCTTAACCCCAAGTAAATTAATTTATTCG 1684986
QY	841	GCTAACTACTTAACCCACGACGCGGCTTCTGTTGCTCCACCATCTCCGAGCAAC 900
DB	1684987	ACCTTACTCTAAACCCAGCTTCTCACGGCGGGCGACCCGTAGCAACAGTATTAAATGAT 1685046
QY	901	GATGCGTTACGTGCGATTGCGGAACAAGAGCTGACTGATATGGCCACGCGTATTACGCGT 960
DB	1685047	GCTCAACTTCGCCAAGATGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1685106
QY	961	ATCGGTCAAGTGTGTTGTTCAATACGCTGCGAGAAAAAGCCCAACCCGACTTCAGCTTT 1020
DB	1685107	ATGCGTCACTTATTCGTTTCAAGTTTAAAAAGAAATATGTTGTCAGAAACAGATTTTCAGCTTT 1685166

Qy 1021 ATCATCAACAGACGCGATGTTCTCTTCACTGCGCCGTGACAAAGAACAAAGTGTCTCGT 1080
Db 1685167 ATCATTGAACAAACCGGTATGTTTCTTTTCACTGCGCCGTGACAAAGAACAAAGTGTCTCGT 1685226
Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTCTCTTCTGTCGCGTAAATGTGGCCGGGATG 1140
Db 1685227 TTAATAAATGAATTTGCGATTTACGCTGTGTTCTGTTCTGTATCAACGTAGCTGGAATC 1685286
Qy 1141 ACACCAGATAACATGCGCTCGCTGCGAAGCGATTTGGCAGTGTCTGTAA 1191
Db 1685287 ACAGAAGATAATTCCTATCTATGTGAAGTATCTGGAAGTACTTTAA 1685337

RESULT 6
US-09-489-039A-4954/c
; Sequence 4954, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 4954
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4954

Query Match 32.5%; Score 387.6; DB 3; Length 579;
Best Local Similarity 79.4%; Pred. No. 2.4e-120;
Matches 459; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 41 TGGGCTGGCGGATCTGTTTCTGTCGATGAACGTCCTCCGCAAAATTAACCTCGGATTC 100
Db 578 TAGTCTGGCCGATCTGTTTCTGTCGATGACCGCTTGAATAATTAACCTCGGATTC 519
Qy 101 GTGCTTATAAAGATGAGACGGGCAAAACCCGGTACTGACGCGTGAAGAAAGGCTGAAC 160
Db 518 GTGTTTACAGGATGAACCGGTAAACGCGCTCTGACGCGTCAAAAGAGCAGAGC 459
Qy 161 AGTATCTGCTGAAATGAACACCAACCAAAATTAACCTCGGATTCGCGGATCCCTGAAT 220
Db 458 AGTATCTGCTGAAATGAACACCAACCAAAATTAACCTCGGATTCGCGGATTCCTGAAT 399
Qy 221 TTGGTCTGCTCAGGAACTGCTGTTTGGTAAAGTACGCGCTGATCAATGACAAAC 280
Db 398 TTGGTCTGCTCAGGAACTGCTGTTTGGTAAAGTACGCGCTGATCAATGACAAAC 339
Qy 281 GTGCTCGACGGCAGACACTCGGGGGGCACTCGGCGCACTACGCGTGGCTGCGGATTTCC 340
Db 338 GCGCCCGCAGCGGCGAGACACCGCGGTAACCGTGGCTGCTGTCGCGCAGACTTC 279
Qy 341 TGGCAAAAATACACGCGTTAAGCGTGTGGGTGAGCAACCCCAAGTGGCGGAAACATA 400
Db 278 TCGCAAAAATACACGCGTTAAGCGTGTGGGTGAGCAACCCCAAGTGGCGGAAACATA 219
Qy 401 AGAGCGCTTTAACTCTGAGGCTCGGAAGTTTGGTAAATACGCTTATATGATGCCGAA 460
Db 218 AAAGCGTATTTCACTCTGCGGGGCTGGAAGTGGCGAATACGATACGATACGAGCGGGCTA 159
Qy 461 ATCAGACTCTTGACTTTGACTTCACTGATTAACAGCGCTGAATGAAGTCAAGCTGGCGAGC 520
Db 158 ACCAGCGGTGACTTTGATGGTCTGCTGGCCAGCTTGAACGAGCCAGCGGGGCGAGC 99
Qy 521 TAGTGTCTTCCATGCTGCTCCATAAACCCCAACCGGTATCGACCTTACGCTGGAAACAAT 580
Db 98 TGTACTGTTCACGGCTGTGCGACAAACCCGACCGGTATCGATCCGACGCTCGATCACT 39

Qy 581 GGCAACACCTGGCAGCAACTCTCGTTTGAGAAAGCTGG 618
Db 38 GGCAGCAGCTGGCGCAGCTCTCGTGGAAAAGGCTGG 1

RESULT 7
US-09-543-681A-3805
; Sequence 3805, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 1999-04-09
; SEQ ID NO 3805
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3805

Query Match 29.3%; Score 348.4; DB 3; Length 702;
Best Local Similarity 70.4%; Pred. No. 5.1e-107;
Matches 466; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTACCGCGCTCTCTGCGACCCGATTTCTGGGCGCTGGCGGATCTGTTT 60
Db 37 ATGTTTGAGAACATTACCGCGCTCTCTGCGACCCGATTTCTGGGCGCTGGCGGATCTGTTT 96
Qy 61 CGTGCCGATGAACGTCCTCCGCGCAAAATTAACCTCGGATTCGTTCTATAAGATGAGAGC 120
Db 97 CGTCTCTGATCTCTGTAACCAAAATTAACCTAGGAGTTCGTTTATAAGGATGAACA 156
Qy 121 GGCAAAACCCGCTACTGACCGCGTGAAGAAAGGCTGAACAGTATCTGCTCGAAATGAA 180
Db 157 GGTAAACACCTGTACTCTACCCCGTTAAAGAAAGCAGAAAAATATTATTAGAAAAAGCA 216
Qy 181 ACCACCAAAATTAACCTCGGCAATTCGCGCATCCCTGAAATTTGCTGCTGCACCTCAGGAA 240
Db 217 TCCACTAAATAATATTCTTCTATTAGGGTATTCTCTGAGTTTGGTGGGTCACTCAAGCA 276
Qy 241 CTGCTGTTTGGTAAAGTACGCGCTGATCAATGAACAAAGTCTGCGCAGGACAGACT 300
Db 277 CTGTTGTTTGGTAAACCAACACCCCTATCATCACAGAAAAACGTCACGTCACGACAGCT 336
Qy 301 CCGGGGGGCGACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACGAGGTT 360
Db 337 CCGGGGGGCGACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCTGCGGACTTTATTGCTCAGCAAACTAATGCT 396
Qy 361 AAGCGTGTCTGGGTGAGCAACCCCAAGCTGGCCGCAACCAATAAGAGCGTCTTTAACTCTGCA 420
Db 397 AAGCGTGTATGATAGTAAAGTAAACCGGCTGGCCCAACCAATAATAATATTATTTCAGACTGCA 456
Qy 421 GGTCTGGAAGTGTGTAATACGCTTATATGATGCGGAAAAATCACACTCTTGACTTCGAT 480
Db 457 GGTCTGGAAGTGTGTAATACGCTTATATGATGCGGAAAAATCACACTCTTGACTTCGAT 516
Qy 481 GCACTGATTAACAGCGCTGAATGAAGCTCAGGCTGGCGAGTAGTGTCTGTTCCATGGCTGC 540
Db 517 GGTATGCTAGTAGTGGCTTACAAAATGCAAGCAGGTGATGTTGTGCTGTTTCCACGAGTGT 576
Qy 541 TGCCTAAACCCCAAGCGTATCGACCTTACGCTGGAAACCAATGGCAAAACACTGGGCACAATC 600
Db 577 TGCCTAAACCCCAAGCGTATGACCCCAAGCTTGAAGCAATGGCGTCAATTTGGCTGCGCTTA 636
Qy 601 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCGAGGTTTTCGCCGTTGT 660
Db 637 TCCGCGAGAAAGGTTGTTTACCGATTTTGAATTTGCTTTACCAAGGTTTTCGTAATGGC 696

QY 661 CT 662
Db 697 CT 698

RESULT 8

US-09-252-991A-12968/c
; Sequence 12968, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12968
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12968

Query Match 26.6%; Score 316.4; DB 3; Length 1509;
Best Local Similarity 54.9%; Pred. No. 5.9e-96;
Matches 645; Conservative 0; Mismatches 526; Indels 3; Gaps 1;

QY 4 TTTGAGAACTTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGCGCGATCTGTTTCGT 63
Db 1290 TTGCCAAGGTGCGCGCGTACCGGCGACCCGATCTCTGGCCCTGCTCGACGCTACCGC 1231

QY 64 GCGGATGAAGTCCCGGCAAAATTAACTCGGAGTTGGTGTCTATAAAGATGAGCGGC 123
Db 1230 AAGCATCCGCGCGGCAAGCTGGACCTCGCGGTGCGGTCTACAAAGGATGCCAGGGC 1171

QY 124 AAAACCCCGTACTGACACGCGTGAAGAGCTGAAACAGTATCTGCTCGAAAAATGAAC 183
Db 1170 CTGACCCCGATCTCGCTCGGTGAACTCGCGGACGCGCTGCTCGAGCAGGAAC 1111

QY 184 ACCAAATTAATCTCGGCATGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGGAATG 243
Db 1110 ACCAAGAGTACGTGCGCGCCACGCGGATGCGCTGTTCCGCGCGCTGCGGGAATG 1051

QY 244 CTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAGTCTCGCAGCGGCACAGACTCG 303
Db 1050 GCGCTCGCGCGCCCTCGCGCTGTTGCTGGAGCAACGCGCCGACCCACGACGCGCC 991

QY 304 GGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGCGAAAAATAACAGCGTTAAG 363
Db 990 GGGGGCACTGGCGCTGCGCTGGCGCGCACTTCACTGCCCATTTGCTGCGCGCGCG 931

QY 364 CGTGTGTGGTAGCAACCCAAAGCTGGCGAACCAATAAGAGCGTCTTTAACTCTGAGGT 423
Db 930 GGCATCTGGCTGAGCGACCCGACCTGGCCGATCCACGAGACCTGTTTCGCGCGCGCG 871

QY 424 CTGGAAGTTCGTGAATAGCTTATTATGATGCGGAAATCACACTTTGACTTCGATGCA 483
Db 870 CTGAAGGTTTCCCACTACCCCTACGTCAGCGCGCGACAAACCGC---CTGGATGTCGAGGCG 814

QY 484 CTGATTAACAGCTTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGCTGCTGTC 543
Db 813 ATGCTTCTGCTGGAGCGCATTTCCCGAGGAGACGTGTGTGCTGCAATGCTGCTGTC 754

QY 544 CATAACCAACCGGTATCGACCTTACGCTGGAACAATGGAACAACTGGCACAACCTCTCC 603
Db 753 CACAACCGACCGGTTTCGACCTGAGCCACGACGACTGGCGAGGGGTGCTCGACGTGGTG 694

QY 604 GTTGAGAAAGGCTGTTACCGCTGTTGACTTACGAGGTTTTCGCCGCTGCTG 663

Db 693 CGTCCGCGGAGCTGCTGCCGCTGATCGACTTTCGCTACGAGGCTTCGCGCAGCGTCTC 634
QY 664 GAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGGCTATGCATAAAGAGCTGATTGTGCC 723
Db 633 GAGGAAGACGCTCGGCGGTACGCTGTTTCGCGCGGAACTGCCGAGGTGCTGGTCAACC 574
QY 724 AGTCTTCTACTCTAAAAACTTTTGGCCTGTACAAAGAGCGTGTGGCGCTTTGACTCTGGT 783
Db 573 AGTTCTCTCGAAGAACTTCGCGCTGTACCGGACCGGCTCGGGCGCTGATCGTCTGC 514
QY 784 GTGCGCGCAGTGAACACCTTGATCGCGCATTCAGCCAAATGAACGCGCGATTCGCGCT 843
Db 513 GCGCAGAACCGCAGAGCTCACCGACCTGCGTAGCCAACTGGCCTTCTCGCCCGAAAC 454
QY 844 AACTACTTAACCCACCGACGACACGCGCTTCTGTTGTGGCCACCATCTCTGACCAACGAT 903
Db 453 CTCTGTGTCGACCCCGCGCGCATGGCGCGAGGTGTCGCGCGCATCTCTCGCGACAGC 394
QY 904 GCGTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATGCGCAGCGTATTCAGCGTATG 963
Db 393 GAGTTGAAGGACTTTGGCAGGAAGAGTCAAGGATCGCTCGCGCATCGCCAGCCTG 334
QY 964 CGTCAAGTTTCTGTCATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTTCAGCTTTATC 1023
Db 333 CGCATCGGCTGTGCGAAGCCCTGCGCGCGCAGCGCTGGCGAGGCTTCGCCCATGTC 274
QY 1024 ATCAAAACAGAACGGCATGTTCTCTTCAGTGGCGCTGACAAAAGAACAGTGTGCGCTCG 1083
Db 273 GCGCGCAACGCGGATGTTTCTCTATACCGACTGAGCCCGCAGCAGGTGCGTCCGGCTG 214
QY 1084 CCGGAGAGTTTGGCGTATATGCGTTCCTGTCGCGTAAATGTCGCCGCGGATGACA 1143
Db 213 CCGCAGAGACGACCGCGTTTACCTGGTGTCCAGCGCGCGGCCAACGTCGCCGCTATAGAC 154
QY 1144 CCAGATAACATCGCTCGCTGTCGAAAGCGATTG 1177
Db 153 GCGCGCGCTCGACCGCTGGCGAAGCCATCG 120

RESULT 9

US-09-252-991A-12507
; Sequence 12507, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12507
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12507

Query Match 26.6%; Score 316.4; DB 3; Length 1632;
Best Local Similarity 54.9%; Pred. No. 5.9e-96;
Matches 645; Conservative 0; Mismatches 526; Indels 3; Gaps 1;

QY 4 TTTGAAACATTACCGCGCTCTCGCGACCCGATTCGGGCTTGGCCGCTGTTTCGT 63
Db 442 TTGCCCAAGGTGCGCGCGGTACCGCGGACCCGATCTTGGGCTGCTCGACGCTACCGC 501
QY 64 GCGGATGAACGTCCCGCGCAAAATTAACCTCGGAGTTGGTGTCTATAAAGATGAGACGGC 123
Db 502 AACGATCCGCGCGCAGCAAGCTGGACCTCGGCGTCCGCTGCTTACAAGGATGCCAGGGC 561

Qy	124	AAAACCCCGGTA	CTGAC	CAGCGT	TGAAAAAGGCT	GAA	CAGTAT	TCTGCT	CGAAAAAT	GAAAC	183
Db	562	CTGACCCCGAT	CTCTCG	CTCGT	TGAAAA	CTCCG	CAGCAG	CGCTGGT	CGAGCAG	GAAAC	621
Qy	184	ACCAAAATTA	CCCTCG	CAAT	TGACGG	CACTCCT	TGAA	TTTGGT	CGCTGC	ACTCAG	243
Db	622	ACCAAGAG	CTACGT	CGCGG	CCAC	CGGCG	ATCG	CGTGT	TCGCG	CGCGCT	681
Qy	244	CTGTTTGGT	AAAGT	TAGCG	CCCTG	ATCA	ATGAC	AAAC	CGTGT	CGC	303
Db	682	CGCTCGG	CGCGCT	CGCCG	CTGT	TG	TGAG	CNA	CGGCG	CAG	741
Qy	304	GGGGCACT	TGCG	CACTA	CGCGT	TGGT	CCGAT	TTTCT	TGGCA	AAAAAT	363
Db	742	GGCGCA	CCGCG	CGCTT	CGCG	CTT	CGCG	CGACTT	CA	TGCGCC	801
Qy	364	CGTGTGT	GGGTG	AGCA	CCCA	CGTGG	CCGA	ACCAT	AAG	ACGTCT	423
Db	802	GGCATCT	TGGT	TGAG	CGAC	CCGAC	CTGG	CCGAT	CC	CAGAG	861
Qy	424	CTGAAAGT	TCGT	GAA	TG	ATAC	GC	GTAT	TAT	GATCG	483
Db	862	CTGAAGT	TTTCCC	ACTAC	CCCTA	CGT	CAG	CGCC	GAC	AACGC	918
Qy	484	CTGATTA	ACAG	CCCTG	AAATG	AGCT	CAG	CGT	CGC	AGCT	543
Db	919	ATGCTT	GTG	GCCT	TGG	AGCG	CA	TTCCC	CAG	GAG	978
Qy	544	CATAA	CCCA	CCG	TAT	TG	ACCT	ACGT	TG	AA	603
Db	979	CACA	CCG	CAG	CCG	TTT	CG	ACT	TAG	CC	1038
Qy	604	GTT	CAG	AAAG	SGCT	GT	TAC	CGCT	GT	TG	663
Db	1039	CGT	CG	CG	AG	CTG	CTG	CC	CT	AT	1098
Qy	664	GAAGA	AGT	GTG	AA	GAG	CT	TCG	CG	CT	723
Db	1099	GAG	GA	AG	CG	CTG	GG	TAC	GC	CT	1158
Qy	724	AGTTC	CTA	CTCT	TAAAA	CTTT	TGG	CTGT	GA	CA	783
Db	1159	AGTTC	CT	CT	CGA	AACT	TCGG	CTGT	AC	CG	1218
Qy	784	GCT	CG	CC	CAG	TGAA	CCG	TAT	CG	CG	843
Db	1219	CGC	CAG	AA	CG	CCG	AG	CT	CG	T	1278
Qy	844	AACT	ACT	CT	TA	AC	CC	CAG	CA	CG	903
Db	1279	CT	CT	GT	CT	GA	CC	CG	CG	CA	1338
Qy	904	CG	CT	AC	T	CG	T	AT	TG	CG	963
Db	1339	GAG	TT	GA	AG	GA	CTTT	GG	CAG	A	1398
Qy	964	CGT	CAG	TT	GT	CT	CA	T	AC	GT	1023
Db	1399	CG	CA	T	CG	CG	CT	GG	CG	CG	1458
Qy	1024	AT	CA	AA	CAG	AA	CG	CA	T	GT	1083
Db	1459	GG	CG	C	CA	CG	CG	GA	T	GT	1518
Qy	1084	CG	GA	AG	AG	TT	TG	CG	GT	T	1143
Db	1519	CG	GA	C	AG	CA	CG	CG	GT	T	1578
Qy	1144	CC	GA	T	AA	C	AT	GG	CT	T	1177
Db	1579	GC	GC	CG	CG	CT	CG	AC	CG	CT	1612

```

RESULT 10
US-08-846-762-1/c
; Sequence 1, Application US/08846762A
; Patent No. 5994072
; GENERAL INFORMATION:
; APPLICANT: Lam, Joseph S.
; APPLICANT: Burrows, Lori
; APPLICANT: Charter, Deborah
; APPLICANT: de Kievit, Teresa
; TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly
; TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa
; FILE REFERENCE: 6580-089
; CURRENT APPLICATION NUMBER: US/08/846,762A
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 24417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-08-846-762-1

```

Query Match	25.7%;	Score	306.6;	DB 2;	Length	24417;			
Best Local Similarity	53.9%;	Pred.	No. 6.9e-92;						
Matches	630;	Conservative	0;	Mismatches	539;	Indels	0;	Gaps	0;
Qy	22	GCTCCTGCCACCCGATTCTGGGCTGGCGATCTGTTTCGTGCGATGAAAGCTCCCGC	81						
Db	23576	GCACCGGTGATCCCATCTTGGCTGAACGAAGCTTCAACGCCGATACCGTTCGGGC	23517						
Qy	82	AAATTTAACTCTGGGATTGGTGTCTATAAGATGAGACGGGCAAACCCCGTACTGACC	141						
Db	23516	AAGATCAACTCGGCTGGCGTGTACTACAACGAGGAGGGCGCATTCGCTGTTCGT	23457						
Qy	142	AGCGTGAANAAGGCTGAAACAGTATCTGCTCGAATAATGAACACCAAAAATTTACCTCGC	201						
Db	23456	GCCGTGCAGGCTGCGAGAAAGCCCGCATCGAGGCCACCGCGCCGCGGTACCTGCGG	23397						
Qy	202	ATTGACGGCATCCCTGAAATTTGGTCGTGCACTCAGGAACTGCTGTTTGGTAAAGGTAGC	261						
Db	23396	ATCGAAGGATCGCCGCCCTACGACCAGGCGGTACAGAACTGCTTTCGGTAAACGAGTCC	23337						
Qy	262	GCCTGATCATGACAAACGTGCTCGACGGCACAGATTCGCGGGGGGCATCGGCCGACTA	321						
Db	23336	GAGCTGTCGGCGCCGCGCTGTGTTGTCACGACCCAGGCGCTCGCGGCACCGCGCGCTC	23277						
Qy	322	CGGCTGGCTCGCGATTTCCTGTCGCAAAAATACACGGTTAAGCGTGTGTTGGGTGAGCAAC	381						
Db	23276	AGCTCGGCGCGCATTTCTCAAGCGCTGTCGCCGACGCCACCGTGGCCATCAGCGAC	23217						
Qy	382	CAAAGCTGGCCGAACCATTAAGAGCGTCTTTAACTCTGCAGGTCTCGAAAGTTCGTAATAC	441						
Db	23216	CCGAGCTGGAAAACACACCGGCACCTGTTTCGAAGCGCGCGCTTCCCGTGCAGAACTAC	23157						
Qy	442	GCTTATTATGATGCGGAAATACACTCTTGACTTCGATGCACTGATTAACAGCCTGAAT	501						
Db	23156	CGCTATTAGACGCGGCCAGCAATGCGGTGAACCCGCGCGGCTGCTCGAAGACCTGAAT	23097						
Qy	502	GAAAGCTCAGGCTGGGACGTAGTGTGTTTCATGGCTGTCGCATAACCCCAACCGGTATC	561						
Db	23096	GCCTCGCGGCGGCTCGATCGTGGTTCTGCAAGCTGCTGCCATAATCCGACGGCGTC	23037						
Qy	562	GACCTCAGCTGGAACAAATGGCAACAATGGCACAACTCTCGTTTGAAGAAAGGCTGTTA	621						
Db	23036	GATCTCGAGCTGGAAGCATGGAACAGGTGTGAGCTGCTCAAGGCCAAGGGCCACGTG	22977						
Qy	622	CGCTGTTTGACTTCGCTTACAGGGTTTTGCCGCTGCTCGAAGAGAGTCTGAAGGA	681						
Db	22976	CGTTCCTCGACATCGCTTACAGGGCTTCGGCAACCGGATCGAAGAAAGCGCCCGCG	22917						
Qy	682	CTCGCGCTTTCGCGGCTATGCAATAAGAGCTGATTGTTGGCAGTTCCTACTCTAAAAAC	741						

Db 22916 GTCCGCTGTTCCGCGCTGGGCTTCTTCTTCCAGCTCGTTCTCCAAATCC 22857
 Qy 742 TTTGGCTGPAACGAGCGTGTGGCGTTGPACTCTCTGTTGCTCCGACAGTGAACCC 801
 Db 22856 TTCTCGCTCTATGGCGAAGCGTGGCGCTCTCGATCGTGACCGAAGCGCGATGAA 22797
 Qy 802 GTTGATCGGCATTCAGCCAAATGAAGCGCGATTTCGGCTAACTACTCTAACCCACCA 861
 Db 22796 TCGGCGCGGTGCTGTCCAGGTGAAGCGGTGATCCGACCAACTATTGCAACCCACCG 22737
 Qy 862 GCACAGCGGCTTCTGTGTGACCATCTCAGCAACGATCGTTAGTGTGCAATTTGG 921
 Db 22736 ACCCAGGCGCAGCTGCTCTCTCGTGCTCAACAGCCCGAACTGCGGGCCCTCTGG 22677
 Qy 922 GAACAAGAGCTGATGATATGCCAGCGATTTCAGCGTATGCGTCAGTTGTTGTCAT 981
 Db 22676 GAGCAGNACTGGCGGAGATGCGGACCGCATCCGCGACATGCGCTGGCGATGTCGAG 22617
 Qy 982 ACCTGAGGAAAGGCGCAACCGGACTTCAGCTTTATCATCAACAGAACGCGATG 1041
 Db 22616 CAACCTGACGCCACGCGCGCAAGCGGACTTCAGTTTGTGTCGTCGCCAACGTGCA 22557
 Qy 1042 TTCTCTTACGTGCGCTGACAAAGAACAAAGTCTGCTGCGCGAAGATTGTCGTA 1101
 Db 22556 TTCTCTTATTCGCGCTGACCGCGACCGAGTTCGAGCGGCTGAAGCCGAGTTTCG 22497
 Qy 1102 TATGCGGTGCTTCTGTCGCTGTAATGTGGCGGATGACACAGATACATGCTGCG 1161
 Db 22496 TACGCGTACGACCGCGCTATCTGCGTGGCGGCTGAACAAATCGAACCTGGAG 22437
 Qy 1162 CTGTGCGAAGCATTTGGCAGTCTGTA 1190
 Db 22436 ATCACAAGGCCATCGTCCAGGTCCTGTA 22408

RESULT 11

US-09-252-991A-12428
 ; Sequence 12428, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12428
 ; LENGTH: 1242
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-12428

Query Match 25.6%; Score 305; DB 3; Length 1242;
 Best Local Similarity 53.8%; Pred. No. 3.6e-92;
 Matches 629; Conservative 0; Mismatches 540; Indels 0; Gaps 0;
 Qy 22 GCTCTGCGACCGGATTTGGGCTGGCGGATGCTGTTTCGTCGCGATGAACGTCCGGC 81
 Db 73 GCACCGGTGATCCATCTTGGCTTGAACGAAGCCCTTCAACGCCGATACCCCTCCGGC 132
 Qy 82 AAATTAACCTCGGATTTGTTCTATAAGATGAGACGGCAAAACCCCGGTACTGACC 141
 Db 133 AAGATCAACCTGGCGTGGCGGTGACTACAAGAGAGGGGCGCATTTCCGCTGTTGCGT 192
 Qy 142 AGCGTAAAGGCTGAAACAGTATCTGCTCGAAATGAAACACCAAAAAATTACCTCGGC 201
 Db 193 GCGTGCAGCTGCCGAGAGGCCGCGCATCGAGGCCACCGCGCGCGGTACTCTGCCG 252

RESULT 12
 US-09-252-991A-12162/c
 ; Sequence 12162, Application US/09252991A
 ; Patent No. 6551795

Qy 202 ATTGACGGCATCCCTGAAATTTGGTGGCTGCACTCAGGAACTGCTGTTTGGTAAAGTAGC 261
 Db 253 ATCGAAGGATCGCGCCCTACACAGGGCGTACAGAACTGCTGTTTGGTAAAGTAGTCC 312
 Qy 262 GGCCTGATCAATGACAAACGTCCTCGACGGCACAGATTCGCGGGGCGACATGCGCGCAT 321
 Db 313 GAGCTGCTGGCGCGCGCGCTGTGTCACGACCGACGGCGCTGCGGGCGACACCGCGCGCTC 372
 Qy 322 CCGTGGCTGCCGATTTCTGSCAAAAATACACGGTTAAAGCTGTGTTGGTGAAGCAAC 381
 Db 373 AAGCTCGCGCGCGACTTCTCCTAAGCGCCTGCTGCGCGACCGCACCGTGGCCATCAGCG 432
 Qy 382 CCAAGCTGCGCGCAACCATAAAGAGCGTCTTAACTCTGCAAGGCTCTGGAAGTTCGTGAAT 441
 Db 433 CCGAGCTGGGAAACCAACCGCGCATGTTTCGAAGCGCGCGCTTCCCGTGCAGAACTAC 492
 Qy 442 GCTTATTATGATCGGAAATACACTCTTGTGATTCGATGCACTGATTAACAGCTTGAAT 501
 Db 493 CGCTATTACGACCGCGCAGCAATGGCGTGAACCGCGCGGCTGCTGGAAGACCTTGAAT 552
 Qy 502 GAAGCTCAGGCTGGCGAGTAGTGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
 Db 553 GGCCTGCGCGCGCTTTCGATCGTGTGCTGCGCGCTGCTGCGCATTAATCCGACCGCGCT 612
 Qy 562 GACCCTAGCTGCAAAATGCGCAACACTCTGCGTTGAGAAAGCTGTTA 621
 Db 613 GATCTGAGCTGAGACACTGGAACAGGTGCTGGAGCTGCTCAAGGCCAAGGCCACGCTG 672
 Qy 622 CGCTGTTTGTACTTCGCTTACAGGGTTTGGCCGCTGCTGGAAGAAAGATGCTGAAGGA 681
 Db 673 CGTCTCTCGACATCGCTTACAGGGCTTCGCGCAACGGCATCGAGGAAGACGCGCGCG 732
 Qy 682 CTGCGGCTTTCGCGGCTATGATTAAGAGCTGATTTGTCAGTTCCTACTCTTAAAC 741
 Db 733 GTGCGCTTGTGCGCGCTGCGGCTGAGCTTCTTCTGTTTCAGCTGCTTCTTCAATCC 792
 Qy 742 TTTGGCTGTACAAAGAGCGTGTGGCGCTTGTACTCTGTTGCTGCGACAGTGAAC 801
 Db 793 TTCTGCTCTATGGCGAAGCGCTGCGCGCCCTCTCGATCGTAACCGAAGCCGCGATGAA 852
 Qy 802 GTTGATCGGCGATTCAGCGCAATGAAGCGCGATTCGCGCTAACTACTCTAACCCACCA 861
 Db 853 TCGGCGCGCTGCTGTCAGGTGAAGCGAGTGATCCGCAACCACTATTTCGAACCCACCG 912
 Qy 862 GCACAGCGCTTCTGTTGTCGCGCATCTGAGCAACGATGCGTTCGTTAGTGGATTGG 921
 Db 913 ACCACGCGCGCGCTGCTCTCTCGCTCAACAGCCCGGAACTGCGGGCCCTCTG 972
 Qy 922 GAACAAGAGCTGACTGATATGCGCGAGCGTATTTCAGCGTATGCTGCTGTTGTTGCTCA 981
 Db 973 GAGCAGAACTGGCGAGATGCGGACCGCATCCGCGCATGCGCTGGCGATGTTGCGAG 1032
 Qy 982 ACGCTGCAGGAAAGGCGCAACCGCGACTTCAGCTTTATCATCAACAGAACGGCATG 1041
 Db 1033 CAACTGGCAGCCACCGCGCGCAAGCGCGACTTCAGCTTTCGTCGTCGCCAACGTGCG 1092
 Qy 1042 TTCTCTTTCAGTGGCTGACAAAAGAAAGAGTGTGCTGCTGCTGCGCGAAGTTTGGCGTA 1101
 Db 1093 TTCTCTTATTCGCGCTGACCGCGCGCGCGCTGAGCGCTGAGCGCTGAGCGCGCTGATC 1152
 Qy 1102 TATGCGTTCCTTCTGCTGCGTAAATGTGGCGGATGACACAGATTAACATGCTGCTCCG 1161
 Db 1153 TAGCGCTGAGCACCGCGCTATTCGCTGCGCGCTGACACAGATTAACATGCTGAGAC 1212
 Qy 1162 CTGTGCGAAGCGATTGTGGCAGTGTGTA 1190
 Db 1213 ATCACAAGGCCATCGTCCAGGTCCTGTA 1241

GENERAL INFORMATION:

```

/ CONTINUATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 12162
/ LENGTH: 1329
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12162

```

```
Query Match      25.6%; Score 305; DB 3; Length 1329;
Best Local Similarity 53.8%; Pred. No. 3.7e-92;
Matches 629; Conservative 0; Mismatches 540; Indels 0; Gaps 0;
```

Qy	22	1212	Db
	GCTCTGCCGACCCGATCTGGGCGCTGGCCGATCTGTTCTGTGCCGATGAACGTCCCGGC	GCACCGGCGTATCCCATCTCTTGGCGCTGAACGAGCGCTTCAACGCGGATACCGTCCGGGC	1153

Qy	82	AAAATTAACCTCGGGATTGGTGCTCTATAAAGATGAGACGGGCAAAACCCCGGTACTGACC	141
Db	1152	AAGATCAACCTTGGGCGGTGGGCGGTGATCTACAACGAGGAGGGGCGCATTCGCTGTTCGGT	1093

	Qy	142	AGCGTGAATAAAGGCTGAACATGTCTGCTCGAAATATGNAACAACCAAAAAATTACCTCCGC	201
	Dd	1092	GCCGTGCAGGTGCCAGAAGGCCCGCATCGAGGCCCAACGCGCGCGGGTGACTCCTGCCG	1033

Qy	202	ATTGACGGCATCCCTGAAATTTGGTCTGCTGCACTCAGGAACTCGTGTTTGGTAAAGGTAGC	261
Db	1032	ATCGAAGGGATCGCCGCTACGACCAGGGCGGTACAGAAACTCGTGTTCGGAAACGAGTCC	973

Qy	262	GCCCTGATCAATGACAAACGTCTCGCACGGCACAGACTCCGGGGGACACTCGCGCACTA	321
Db	972	GAGCTGCTGGCGCGCGCCGTGTGTCACGACCCAGCGCCCTCGCGGGCACGCGCGCGCTC	913

QY 322 CGCGTGGCTGCCGATTTCTGGCAAAATACCAAGCGTTAAGCGTGTGTGGTGAGCAAC 381
DB 912 AAGCTCGGCGCGGACTTCCTCAAGCGCTGTGTCGCCAAGCCACCGTGGCCATCAGCGAC 853

Qy	382	CCAAGCTGGCGCAACCATTAAGAGCGTCTTTAACTCTGCAGGCTTGGAAGTTTCGTGAATAC
Db	852	CCGAGCTGGGAAAACCACCGGCACGTTTCGAAAGCCGCCGTTTCCCGGTGCAGAACTAC

Qy		442
GCTTATTATGATCGGGAAATACACACTCTTGACTTCGATGCACTGATTAAAGCCTGAAT		501
Db		792
GCCTATTACGACGCCCGCCAGCAATGGCTGTAAACCGCCGGCTCTGTGGAGAACCTGAAT		733

QY 502 GAAGCTCAGGCTGGCGACGTAGTCTGTTCATGGCTGCTGCATACCCAAACCGGTATC 561
 Db 732 GCCCTGCGGGCCGTTTCGATCGTGGTCCTGCACGCTGCTGCATAATCCGACCGCGTC 673

QY 562 GACCTACGCTGGAAACAATGGCAAACTGGCACAACTCTCGTTGAGAAACGCTGGTTA 621
 |||||
 Db 672 GATCTCGAGCTGGACGACTGGAAAACAGGTGCTGGACGTCTCAAGGCCAAGGGCCACGTG 613

QY	622	CGCGTGTGACTCTGCTTACAGGGGTTGCCCCGGTCTCGAAGAAGATGCTGAAGGA	681
Db	612	CCGTTCTCGACATCGCTACAGGGGTTTCGGCAACGGCATCGAGGAAGACGCCCGCGCG	553

QY
682
552

DB

741
493

QY 11GGCCGTATCAACAGACGGTGTGGCCCTGTGATCTGGTCTGCCGACAGTGAACC 801
 742 11GGCCGTATCAACAGACGGTGTGGCCCTGTGATCTGGTCTGCCGACAGTGAACC 801
 Db 492 TTCTCGCTCTATGGCGAACGGCTCGGCGCCCTCTCGATCGTAACCGAAGCGCGCATGAA 433

Qy	802	GTTGATCGCGCATTTACGCCAAATGAAGAGCGCGGATTCGCGGTAACTACTCTAAACCCACCA	861
Db	432	TCGCGCCGCGTGCTGTCCAGGTGAAGCGAGTGTATCCGACCAACTATTTCGAACCCACCG	373
Qy	862	GCACACGGCGCTTCTGTGTGTTGCCACCATCTGTAGCAACGATCGGTACGTCGCAATTGG	921
Db	372	ACCCACGGCGCCAGCGTCGTCTCTTCGTGCTCAACAGCCCGGAACCTCGGCGCCCTCTGG	313
Qy	922	GAAACAAGAGCTGACTGATATATCGCGCAGCGTATTGAGCGTATGCGTCAGTTGTTCTGTCAA	981
Db	312	GAGCAGGAATCTGGCGAGATCGCGCAGCGCATCCGCGACATGCGCTTGGCGATGTCGAG	253
Qy	982	ACGCTGCAGGAAAAAGGCGCAAAACCGGCACTTCAGCTTTATCATCAAAACAGAACCGCATG	104
Db	252	CAACTGCGACGCCACGGCGCCCAAGCGCGACTTCAGCTTCGTTCGTTCGCGCAACGTCGGCATG	193
Qy	1042	TTCTCCTTCAGTGCCTTCACAAAAGAACAACTGCTGCGCGAAGAGTTTCGGCGTA	110
Db	192	TTCTCCTATTTCGCGCCTGACCGCGCACGAGTCGAGCGCCCTGAAGACCGAGTTCGGTATC	133
Qy	1102	TATCGGTTGCTTCTGCTCGCGTAAATGTGCGCGGGATGACACAGATAAATGCTCTCG	116
Db	132	TACGCGTCAGCACCGGCGGTATCTGCGTCGCGCGCTGAACAAATCGAACTGGAGACC	73
Qy	1162	CTGTGCGAAGCGATTGTGGCAGTCTGTA	1190
Db	72	ATCACCAAAGGCCATCGTCCAGGTCCTGTA	44

RESULT 13

```

US-09-489-039A-1301
; Sequence 1301, Application US/09489039A
; Patent NO. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1301
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1301

```

Query Match 24.3%; Score 289; DB 3; Length 1338;
Best Local Similarity 53.2%; Pred. No. 9.7e-87;
Matches 635; Conservative 0; Mismatches 555; Indels 3; Gaps 1

Qy 2 TGTTTGAGAACATTACCGCGCTCTCTGCGACCGAATTCCTGGCGCTGGCGGATCTGTTTC 61

Db 146 TGTTCCTAAAGATTGACGCTACGCGCGGACCTATTCTTTCTGTATGGACGCTTCA 205

Qy 62 GTGCCGATGAACGTCCCGGCMAAATTACCTCGGATTTGGTGTCTATAAAGATGACCG 121

Db 206 AAGAAGACCCCGGAAGCGACAAAGTCAACCTGAGTATCGGGCTGTACTACACGATGAC 265

Qy 122 GC AAAACCCCGGTACTGAC CAGCGTGAAAAAGCTGAACAGTATCTGCTGCAAAATG --- 178

Db 266 GCATTATCCCGCAGCTG CAGCGCGGTGGCGGAAGCGGAACACCGCTGAACCCGAGCCGC 325

Qy 179 AAACCCACAAAAATTACTCTGGGCAATTGACGGCATCCCTGAAATTTGGTTCGTGCACTCAGG 238

Db 326 ATGGCGCCTCGTGTAATCTGCCAATGGAAGGGTTGAGCGGCTACCGTCAGGCGGATTTGGC 385

Qy
239 AACTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCACGGCACAGA 299
Db
386 CACTGCTGTTTGGCGCAGAGCATACCGCGCTTAAGCAAAATCGCATTCGCTCTATTTCAGA 445

Qy 899 ACGATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCGCAGGTATTCAGC 958
 Db 239 ATGCGGGCTGAAGCCACGTCGGCAGCGGGAAGTGGAGCGCATGCGCGCGCACATCCTGA 180
 Qy 959 GTATGCGTCAGTTGTTGTCATACGTCGACAGGAAAGGCGCAACCGCGACTTCAGCT 1018
 Db 179 CCATGCGCAGCGCTGTCGACGCGCTGCGAGAGGTTGCCCGGCGAGCAAGTGGATT 120
 Qy 1019 TTATCATCAACAGCAGCGCATGTTCTCTTTCAGTGGCTGACAAAAGAAACAAGTGGCTG 1078
 Db 119 ACCTGCTCAAGCAGCGCGGATGTTACGTATACCGGATTTAGCGCGCGCAGTGGATC 60
 Qy 1079 GTCTGCGGAGAGTTTGGCGTATATGCGGTGCTTCTGTGCGGTGAAATGTCGCCGG 1137
 Db 59 GTCTGCGCATAGTTTGGCGTCTACCTGATTGCGAGCGGACGATGTCGCGTGGCGGG 1

RESULT 15
 US-09-540-236-315
 ; Sequence 315, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
 ; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540.236
 ; CURRENT FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 315
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: M.catarrhalis
 US-09-540-236-315

Query Match 23.2%; Score 276.4; DB 3; Length 1209;
 Best Local Similarity 52.6%; Pred. No. 1.7e-82;
 Matches 626; Conservative 0; Mismatches 561; Indels 3; Gaps 1;

Qy 1 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTTCTGGCCCTGGCGCTCTGTTT 60
 Db 13 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTTCTGGCCCTGGCGCTCTGTTT 72
 Qy 61 CGTGCCGATGAACGTCCTCGGC---AAAATTAACTCGGATTTGGTGTCTATAAGATGAG 117
 Db 73 GCCAAGACCCAGCATGACATTAAGTCAATCTTGGTGTGGCGGTGACTATACCGAA 132
 Qy 118 ACGGGCAAAACCCGGTACTGACAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAT 177
 Db 133 GATGGTCAATTGCTGTACTTGAATGCTCAAAACCCCGAATCCAAAATCACCATCCA 192
 Qy 178 GAAACCAACCAAAATTAACCTCGGATGACGGCATCCCTGNAATTTGGTGTGCTGACTCAG 237
 Db 193 CCTGCTCCAGGAGATATTTGCCAATGGATGGTTTGGCAGGTTATCGTAAGGCGATGCCAA 252
 Qy 238 GAATGCTGTTTGGTAAAGGTAGCGCTGATCAATCAACAAACGCTGCTCGCAGCGCACAG 297
 Db 253 GATTTGTTATTTGGTAAAAATCATCTGCGAGTTTATAGGGGCGTGTGCCACGATTGCA 312
 Qy 298 ACTCCGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATPACCAGC 357
 Db 313 ACTTTGGGGGTTCTGCTGCGTTAAAGGTTGGTGTGCTGATTTTATCCATGAGTGGTTTCT 372
 Qy 358 GTTAAGCGTGTGGGTGAGCAACCCAGCTGGCGGAACCATAGAGCGTCTTTAACTCT 417
 Db 373 AATGCCAAATGCTATGTGTGAGACCCAACTTGGGCAAAATCATGTGGGTATTTTGGGGT 432
 Qy 418 GCAGGCTGGAAGTTCGTGAATACGCTTATTATGATCGGGAATCACACTCTTGACTTC 477
 Db 433 GCAGGTTTGGGTGGCAAGTATCTCTTATATGATCCAAAACATATTGGGCTAAAAATTT 492
 Qy 478 GATGCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGC 537

Db 493 GATGAGCTGTGCGAATTTTTCAAGACTTTTAAATGAAATGATGTAGTGTACTACACCA 552
 Qy 538 TGTGCCCATAACCCAAACCGGTATCGACCTACGCTGGAACAATGGCAAACTCTGGCAAA 597
 Db 553 TGTGCCCATAACCCAACTGTGTGTGATTTGACCCGTGAACAGTGGAAATGTACTATTACAA 612
 Qy 598 CTCTCCCTTGAGAAAGGCTGTACCGCTGTTCACCTTTCGCTTACCTTACAGGGTTTGGCCGT 657
 Db 613 ATCGTCAAGATATAAAAACTTATTCATTTATGATATTCCTATCAAGGTTTGGCGAT 672
 Qy 658 GGTCTGGAAGATGCTGAAGACTGCGCGCTTTCGCGGCTATGTCATAAAGAGCTGATT 717
 Db 673 GATATGATGGTGTGATGCTATGCGATTCGTCGCGGTGGGGCTGGGATTCACGCTATTT 732
 Qy 718 GTTGCCAGTTCTTACTCTAAAACTTTGGCCTGTACAAACGAGCGTGTGGCGCTTGTACT 777
 Db 733 GTCTCTAATTCATTTTCAAAAACTTGTCTACTTTATGGTGAGCGTGTGGCGGCTTATCT 792
 Qy 778 CTGCTGTGCTCCGACAGTGAAACCGTTGATGCGGCATTCAGCCAAATGAAAGCGCGATT 837
 Db 793 GTCGTTGACCAACCAAGAAAGAGAGATGTGTGTAATTCGTCATTAATAATTTACCGTC 852
 Qy 838 CGGCTTAATCTACTTAACCCACGACGACGCGCTTCTGTGTGTCACCATCTGAGC 897
 Db 853 CGTCGCAATTTATCAAGTCCAGCTGCACGCGCAACTATGTGTGTGATGCGGTCAATGAAT 912
 Qy 898 AACGATGCGTTACGTGCGATTTGGGAACAAGAGCTGATGATATGCGCGCAGCTATTTCAG 957
 Db 913 GATGAGCATTTATTTGGCTTATGGATTGGTGAGGCATATGAAATGCGTGACCGCATTCGT 972
 Qy 958 CGTATGCGTCAGTTGTTGCTCAATAACGCTGACGGAAGGCGCAACCCGCGACTTCAGC 1017
 Db 973 GATATGCGTCAAAAACTCCAAGAGCATTAAGCACTAAACTGCTGAGCGTGATTTTATGC 1032
 Qy 1018 TTTTATCATCAACAGAACGCGCATGTTCTCTTCTGTCGCTGACAAAGAAACAAGTGTCTG 1077
 Db 1033 TATTTTACCAACAGCGCTGTGTATGTTTAGCTTTTACAGGACTTAACCAAGAGCAAGTGTATT 1092
 Qy 1078 CGTCTGCGCAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTGGCCGG 1137
 Db 1093 CGTTTGGGTGAAGATTTGGGTATATATGTTGAAATGGCGTATGTGTGTGGCAGGT 1152
 Qy 1138 ATGACACAGATTAACATGCTCCGCTGTGGAAGCGATTGTCGAGTGCT 1187
 Db 1153 CTGAATAATTTCTAATGTAGAATATGTTGCCAGTGCTATGCTGAAGTGCT 1202

Search completed: March 15, 2006, 03:10:27
 Job time : 459 secs

This Page Blank (uspto)

GenCore version 5.1.7.
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:26:24 ; Search time 3885 Seconds
(without alignments)
4769.031 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAAPADPILGLDLF.....VAGTMDNMAPLCEAIVAVL 396

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/abss/ABSSWEB_spo01/US10673786/runat_14032006_135611_22627/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss05p
-USER=US10673786 @CGN 1.1.8010 @runat_14032006_135611_22627 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1486	72.7	864	10	CL660432
2	1352	66.1	785	10	CL695845
3	989	48.4	645	6	CA0933007
4	947	46.3	671	10	CL695172
5	914	44.7	1690	4	AY103658
6	909	44.4	1749	8	CX109239
7	868	42.4	1599	4	CNS0A2EC

8	857	41.9	1832	4	AY105250
9	835	40.8	557	6	CF306417
10	835	40.8	1486	4	CNS092WQ
11	807.5	39.5	1502	4	AY815146
12	797	39.0	2452	4	CR857622
13	795	38.9	468	1	AW221110
14	777	38.0	2447	4	CR860094
15	759	37.1	1496	4	CR616132
16	759	37.1	1509	4	CR626608
17	759	37.1	1512	4	CR605482
18	759	37.1	1525	4	CR615083
19	759	37.1	1540	4	CR598904
20	759	37.1	1560	4	CR622320
21	759	37.1	1563	4	CR601723
22	759	37.1	1719	4	CR606661
23	759	37.1	1959	4	CR625122
24	759	37.1	1979	4	CR619243
25	759	37.1	1979	4	CR620721
26	759	37.1	2017	4	CR860603
27	754	36.9	1564	4	CNS0EPVU
28	748	36.6	1517	4	CNS0GMY6
29	746.5	36.5	1510	4	AY223393
30	745	36.4	1440	4	CNS0EOL3
31	739.5	36.2	2019	4	CR926063
32	736	36.0	1643	4	CNS0A80C
33	734	35.9	2050	4	CR861057
34	733.5	35.9	977	7	CK259742
35	732	35.8	968	7	CK262759
36	721	35.3	1704	4	CR625110
37	718	35.1	1566	4	CNS0A3OI
38	718	35.1	1762	4	CNS0GRLY
39	712	34.8	1771	4	CNS0EYV2
40	709	34.7	1034	7	CV734935
41	708	34.6	1358	7	CO635670
42	706.5	34.5	902	7	CK259741
43	705	34.5	902	7	COL13549
44	702	34.3	1131	7	CV734832
45	697	34.1	1102	10	AY407844

ALIGNMENTS

CL660432 864 bp DNA linear GSS 09-JUL-2004
PRI0137a_D10 - PRI0137a.B21 (864) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL660432
CL660432.1 GI:50145720
GSS
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 864)
Srinivasan, J., Otto, G.W., Kahlow U., Geisler, R. and Sommer, R.J.
AppaDB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: fosmid ends.

Location/Qualifiers

```

source
1. .864
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Alignment Scores:
Pred. No.: 1.36e-160 Length: 864
Score: 1486.00 Matches: 285
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 72.7% Indels: 0
DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x CL660432 (1-864)
Qy 19 LeuPheArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyLysAsp 38
Db 10 CTGTTTCGCGGATGAACGTCGCCGCAAAATTACCTCGGATTTGGTCTTAAAGAT 69
Qy 39 GluThrGlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLysLeuGlu 58
Db 70 GAGACGGCAAAACCCCGGTACTGACGAGCGTGAAGAGGCTGAACGATATCTGCTCGAA 129
Qy 59 AenGluThrThrLysAenTyLysLeuGlyIleAspGlyIleProGluPheGlyArgCysThr 78
Db 130 AATGAACCCACCAAAATTTACTCTGGCATTACCGCATCCCTGGAATTTGGTCGTGCAC 189
Qy 79 GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAla 98
Db 190 CAGGAACCTGCTTTGGTAAAGTAGAGCCCTGATCAATGACAAACGTCGTCGACGGCA 249
Qy 99 GlnThrProGlyGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThr 118
Db 250 CAGACTCGGGGGGCACTGGCGCACTACGCTGGCTGCCGATTTCTTGGCAAAATAACC 309
Qy 119 SerValLysArgValTrrValSerAenProSerTrpProAenHisLysSerValPheAen 138
Db 310 AGCGTTAAGCGTGTGGGTGAGCAACCAACAGCTGGCGCAACCAATAGAGCGCTCTTAAC 369
Qy 139 SerAlaGlyLeuGluValArgGluTyLysAlaTyLysAspAlaGluAenHisThrLeuAsp 158
Db 370 TCTGAGGCTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAATAATCACACTCTTGAC 429
Qy 159 PheAspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHis 178
Db 430 TTCGATGCACCTGATTTAACAGCGTGAATGAAGCTCAGGCTGGCGAGTAGTCTGTTCAT 489
Qy 179 GlyCysCysHisAenProThrGlyIleAspProThrLeuGluGlnTrrPheLeuAla 198
Db 490 GGCTGCTGCCATAACCCCAACCGGTATGACCCCTACGCTGGCAACATGCGCAACACTGGCA 549
Qy 199 GlnLeuSerValGluLysGlyTrrLeuPheAspPheAlaTyLysGlnGlyPheAla 218
Db 550 CAACTCTCCGTTGAGAAAGCGTGGTTACCGCTGTTTGTGACTTCGCTTACAGGGTTTGCC 609
Qy 219 ArgGlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
Db 610 CGTGGCTGGAAGAAGATGCTGAAGGAGCTGCGCGCTTTTCGCGCTATCATGAAGAGCTG 669
Qy 239 IleValAlaSerSerTySerLysAenPheGlyLeuTyLysAenGluArgValGlyAlaCys 258
Db 670 ATGTTGTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 729
Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
Db 730 ACTCTGGTTGTCGCCGACAGTGAACCGTTGATCGCGCATTCAGGCCAATGAAAGCGCG 789
Qy 279 IleArgAlaAenTyLysSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeu 298

```

```

Db 790 ATTGCGCTAACTACTCTAAACCCAGCAGCACGGCGCTTCTGTGTGCCCATCTG 849
Qy 299 SerAenAspAlaLeu 303
Db 850 AGCAACGATCGTTA 864

RESULT 2
CL695845 785 bp DNA linear GSS 10-JUL-2004
LOCUS PRI017b.H12.2 - PRI017b.BR (785) Mixed stage fosmid library of P.
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL695845
VERSION CL695845.1 GI:50217753
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Rukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 785)
AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppaDB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED 14681447
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
Location/Qualifiers
source
1. .785
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Alignment Scores:
Pred. No.: 3.63e-145 Length: 785
Score: 1352.00 Matches: 260
Percent Similarity: 99.6% Conservative: 0
Best Local Similarity: 99.6% Mismatches: 1
Query Match: 66.1% Indels: 0
DB: 10 Gaps: 0

US-10-673-786A-2 (1-396) x CL695845 (1-785)
Qy 128 ProSerTrpProAenHisLysSerValPheAenSerAlaGlyLeuGluValArgGluTyL 147
Db 3 CCAAGCTGGCGGCAACCAATAGAGCGCTTTTAACCTCTGAGGCTCTGGAAGTTCTGTGATAC 62
Qy 148 AlaTyLysAspAlaGluAenHisThrLeuAspPheAspAlaLeuIleAenSerLeuAen 167
Db 63 GCTTATTATGATCGGGAATACACACTCTTGACTTCGATCGACTGATTAAACAGCTGAAT 122
Qy 168 GluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCysHisAenProThrGlyIle 187
Db 123 GAAGCTCAGGCTGGCGAGTAGTGTCTTCATCGCTGCTGCCATAACCCCAACCGGTATC 182
Qy 188 AspProThrLeuGluGlnTrrPheLeuAlaGlnLeuSerValGluLysGlyTrrLeu 207
Db 183 GACCTTACGCTGGCAACATGGCAAACTGGCACAACTCTCCGTTGAGAAAGGCTGGTTA 242

```


QY 208 ProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeuGluGluAspAlaGluGly 227
 Db 243 CGCGCTTTGACTTCTTACAGGGTTTGGCCGGTGTCTGGAAGAAGATGCTGAAGA 302
 QY 228 LeuAtqAlaPheAlaMetHisLysGlyLeuLeuValAlaSerSerTyrSerLysAsn 247
 Db 303 CTGCGCGCTTTGCGGCTATGATATAAGAGCTGATTGTTGCCAGTTCCTACTCTAAAAAC 362
 QY 248 PheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuValAlaAlaAspSerGluThr 267
 Db 363 TTGGCTCTACACAGAGCGTGTGGCGCTTGTACTCTGGTGTCTGCCAGACAGTGAACC 422
 QY 268 ValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAlaAsnTyrSerAsnProPro 287
 Db 423 GTTGATCGCGCATTCAGCCAAATGAAAGCGGCGATTGCGGCTAACTACTCTAACCCACCA 482
 QY 288 AlaHisGlyValSerValAlaThrIleLeuSerAsnAspAlaLeuArgAlaIleTTP 307
 Db 483 GCACAGCGCGCTTCTGTTGTCACCCATCTGAGCAACGATGCGTTAGTCGATTTGG 542
 QY 308 GluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValAsn 327
 Db 543 GAACAGAGCTACTGATATGCGCCAGCGTATTACAGCGTATGCGTCAGTTGTGTCAT 602
 QY 328 ThrLeuGlnGluLysGlyAlaAsnArgAspPheIleIleLysGlnAsnGlyMet 347
 Db 603 ACGTGCAGGAAAGGCGCAACCGCGACTTTCAGCTTTATCATCAACAGAGCGCATG 662
 QY 348 PheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeuArgGluGluPheGlyVal 367
 Db 663 TTCTCCTTCAGTGGCGCTGACAAAGAACAAAGTGTGCGTCTGCGCGAAGAGTTTGGCGTA 722
 QY 368 TyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThrProAspAsnMetAlaPro 387
 Db 723 TATGCGGTGCTTCTGTCGCTAAATGTGGCGGATGACACAGATAACATGTTCCG 782
 QY 388 Leu 388
 Db 783 CTG 785

RESULT 3
 CA093007
 LOCUS
 DEFINITION 645 bp mRNA linear EST 23-SEP-2003
 3', mRNA sequence.
 CA093007
 CA093007.1 GI:34946314
 EST.
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 1 (bases 1 to 645)
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 The libraries that made SUCST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 002 row: F column: 08
 Seq primer: M13/Forward primer.
 Location/Qualifiers
 1. .645

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source

/organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCCCL1002F08"
 /lab_host="XLI Blue MRP"
 /clone_lib="CL1"
 /note="Organ: Pool of sugarcane calli submitted to low
 (40C) and high (37 C) temperature stress; Vector:
 pBluscript; Site 1: EcoRI; Site 2: XhoI; An
 unidirectional cDNA library generated from [Pool of
 sugarcane calli submitted to low (40C) and high (37 C)
 temperature stress]. cDNA was prepared from polyA+ mRNA
 using ZAP - cDNA Synthesis kit (Stratagene). The
 double-strand cDNAs were fractionated in a sphaerose CL-2B
 40cm-columns and fragments sizing between 0.8 and 1.5 Kb
 were directionally cloned into the vector. Details of
 each source of RNA and library construction can be
 obtained at http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
 Pred. No.: 2,47e-103 Length: 645
 Score: 989.00 Matches: 186
 Percent Similarity: 97.9% Conservatives: 1
 Best Local Similarity: 97.4% Mismatches: 4
 Query Match: 48.4% Indels: 0
 DB: 6 Gaps: 0
 US-10-673-786A-2 (1-396) x CA093007 (1-645)

QY 31 LeuGlyIleGlyValTyrLysAspGluThrGlyLysThrProValLeuThrSerValLys 50
 Db 22 CTCGGATTGGTGTCTATAAAGATGAGACGGGCAAAACCCCGTACTGACGCGTGAAA 81
 QY 51 LysAlaGluGlnTyrLeuLeuGluAsnGluThrThrLysAsnTyrLeuGlyIleAspGly 70
 Db 82 AAGCGTGAACAGTATCTGCTCGAAAATGAAACACCAAAATTTACCTCGGCATGACGGC 141
 QY 71 IleProGluPheGlyArgCysThrGlnGluLeuLeuPheGlyLysGlySerAlaLeuIle 90
 Db 142 ATCCCTGAAATTGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 201
 QY 91 AsnAspLysArgAlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuArgValAla 110
 Db 202 AATGACAAACGCTCTCGACGCGACAGACTCCGGGGGCGACCTGCGGCACCTACGCGT 261
 QY 111 AlaAspPheLeuAlaLysAsnThrSerValLysArgValTyrValSerAsnProSerTTP 130
 Db 262 GCCGATTTCTCGGCAAAATAACACGCGTTAAGCGTGTGGGTGAGCAACCCAAAGCTGG 321
 QY 131 ProAsnHisLysSerValPheAsnSerAlaGlyLeuGluValArgGluTyrAlaTyrTyr 150
 Db 322 CCGAACCATTAAGAGCGTCTTTAACTCTGACAGTCTGGAAGTTCTGTAATACGCTTATTAT 381
 QY 151 AspAlaGluAsnHisThrIleuAspPheAspAlaLeuIleAsnSerIleuAsnGluAlaGln 170
 Db 382 GATCGGAAAAATCACACTCTTGACTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
 QY 171 AlaGlyAspValValLeuPheHisGlyCysCysHisAsnProThrGlyIleAspProThr 190
 Db 442 GCTGGCGACGTAGTGTCTGTTCCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
 QY 191 LeuGluGlnTTPGlnThrLeuAlaGlnLeuSerValGluLysGlyTTPLeuProLeuPhe 210
 Db 502 CTGGAACAAATGGCAAAACACTGGCAACTCTCCGTTGAGAAGGCTGTTTACCGCTGTTT 561
 QY 211 AspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
 Db 562 GACTTCGCTTACCAAGGTTGCCCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594

RESULT 4
 CL695172
 LOCUS

CL695172 671 bp DNA linear GSS 10-JUL-2004

```

DEFINITION PR10166a_G11.2 - PR10166a.BR (671) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic survey
            sequence.
ACCESSION  CL695172
VERSION     CL695172.1  GI:50217080
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
ORGANISM    Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 671)
AUTHORS     Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE       AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
JOURNAL     Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED      14681447
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@uebingen.mpg.de
            This library was generated at Caitech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES    Location/Qualifiers
            1..671
               /organism="Pristionchus pacificus"
               /mol_type="genomic DNA"
               /strains="California"
               /db_xref="taxon:54126"
               /clone_lib="Mixed stage fosmid library of P. pacificus
               var. California"
               /note="Vector: pBpifos-5 Fosmid vector"
ORIGIN
Alignment Scores:
Pred. No.:      1,88e-98      Length:      671
Score:          947.00      Matches:    190
Percent Similarity: 97.9%      Conservative: 0
Best Local Similarity: 97.9%      Mismatches: 3
Query Match:    46.3%      Indels:      1
DB:             10      Gaps:          0

US-10-673-786A-2 (1-396) x CL695172 (1-671)
QY 204 LysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeuGluGlu 223
    |||||
Db 1  AAAGGCTGGTTACCGCTGTTGACTTCGCTTACCAGGGGTTTGGCCGGTGGTCTGGAGAA 60
QY 224 AspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleValAlaSerSer 243
    |||||
Db 61 GATGCTGAAGGACTGCGCGCTTCGCGCTATGCATAAAGAGCTGATTGTTGCCAGTTCC 120
QY 244 TyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuValAlaAla 263
    |||||
Db 121 TACTCTAAAAACTTTGGCGCTGTACACGAGCGTGTGGCGCTGTGACTCTCGTTGCTGCC 180
QY 264 AspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAlaAsnTyr 283
    |||||
Db 181 GACAGTGAACCGTTGATCGCGCTTACGCAATGAAGCGCGGATTCGCGCTAACTAC 240
QY 284 SerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAspAlaLeu 303
    |||||
Db 241 TCTAACCCACGACACACGCGCTCTCTGTTGTGTCACCATCTCTGAGCAACGATCGGTTA 300
QY 304 ArgAlaIleTyrGlu-GlnGluLeuThrAspMetArgGlnArgIleGlnArgMetArgG1 323
    |||||
Db 301 CGCGCCGATTGGGAAACACGAGAGCTGACTGATATGCGCCAGCGTATTACGCGTATGCGTCA 360
QY 323 nLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPhePheIleLeLy 343

```

```

Db 361 GTTGTTCGTAATACGCTGAGGAAAAGGCGCAACCGGACTTCAGCTTATCATCAA 420
QY 343 sGlnAsnGlyMetPheSerGlyLeuThrLysGluGlnValLeuArgLeuArgG1 363
    |||||
Db 421 ACAGAAGCGCATGTTCTCTTCCTTCAGTGGCCTGACAAAAGAACAAAGTGTGCGTCTGCCGA 480
QY 363 uGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThrProAs 383
    |||||
Db 481 AGAGTTTGGCGTATATCGGTTGCTTCTGCTCGGTAATGTCGCGGATGACACCAGA 540
QY 383 pAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
    |||||
Db 541 TAACATGTTCCGCTGTGCGAAGCATTTGTGGCAGTGCTG 580

RESULT 5
LOCUS      AY103658                1690 bp      mRNA      linear      HTC 18-FEB-2005
DEFINITION Zea mays PCO118382 mRNA sequence.
ACCESSION  AY103658
VERSION     AY103658.1  GI:21206736
KEYWORDS    HTC.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 1690)
            Gardiner,J., Schroeder,S., Polacco M.L., Sanchez-Villeda,H.,
            Fang,Z., Morgante,M., Landewe,T., Fengler,K., Useche,F.,
            Hanafey,M., Tingey,S., Chou,H., Wing,R., Soderlund,C. and Coe,E.H.
            Jr.
            Anchoring 9,371 maize expressed sequence tagged unigenes to the
            bacterial artificial chromosome contig map by two-dimensional
            overgo hybridization
            Plant Physiol. 134 (4), 1317-1326 (2004)
JOURNAL     PUBMED
            15020742
REFERENCE   2 (bases 1 to 1690)
AUTHORS     Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
            Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
            Unpublished (2002)
REFERENCE   3 (bases 1 to 1690)
AUTHORS     Coe,E.H.
            Direct Submission
            Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
            If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES    Location/Qualifiers
            1..1690
               /organism="Zea mays"
               /mol_type="mRNA"
               /db_xref="maizemap:636895"
               /db_xref="taxon:4577"
               /clone_lib="Maize Mapping Project/DuPont Consensus
               Library"
               /note="this sequence is part of a project of EST
               assemblies resulting from the application of public
               contigs to seed DuPont contigs; this resource was
               assembled by DuPont as part of a collaboration for the
               overgo addressing of BACs in conjunction with the Maize
               Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.:      5e-94      Length:    1690

```

Db	1203	ATTAACATGAGGCCAACCACTATTATAATGCCTGAATCCAGAGAACCCTCGTGATGG	1262
Qy	339	SerPheIlellellyGsInAnGlyMetPheSerPheSerGlyLeuThrLyGsGluGlnVal	358
Db	1263	AGCATATCATTANGCAAATTTGGATGTTTACTTCTTCTGCTGTGATACCGAACAAGTC	1322
Qy	359	LeuArgLeuArgGluGluPhueGlyValTyAlaValalaSergGlyArgValanValala	378
Db	1323	GCAATTCATGAGGAGGAATACCACTTATATGACATCTCATGGAGGATCAGCATGGCC	1382
Qy	379	GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValalavalLeu	396
Db	1383	GGTTTGATGACATGAGGACTGTGCCCCACTTTTCAGATGCCATACACGCTGCAGTT	1436
 RESULT 6 CX109239			
LOCUS	CX109239	1749 bp	mRNA linear EST 03-JUN-2005
DEFINITION	RECOM798 A normalized whole-life-cycle cdna library of rice Oryza sativa (indica cultivar-group) cDNA clone BI130A13, EI078H06, BI100J15, BI107E19, BI131F09, BI075E21, BI07 5', mRNA sequence.		
ACCESSION	CX109239		
VERSION	CX109239.1 GI:66922391		
KEYWORDS	EST.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriartoideae; Oryzaceae; Oryza. .1 (bases 1 to 1749)		
REFERENCE	Zhang,J., Feng,Q., Jin,C., Qiu,D., Zhang,L., Xie,K., Yuan,D., Han,B., Zhang,Q. and Wang,S. Features of the expressed sequences revealed by a large-scale analysis of ESTs from a normalized cDNA library of the elite indica rice cultivar Minghui 63		
JOURNAL	Plant J. 42 (5), 772-780 (2005)		
PUBMED	15918889		
COMMENT	Contact: wang S National Key Laboratory of Crop Genetic Improvement Huazhong Agricultural University Wuhan 430070, China Tel: 86-27-87282044 Fax: 86-27-87287092 Email: shipingwang@hotmail.com Seq primer: T7.		
FEATURES	Location/Qualifiers		
source	1..1749 /organism="Oryza sativa (indica cultivar-group)" /mol_type="mRNA" /strain="indica" /cultivar="Minghui 63" /db_xref="taxon:39946" /clone="BI130A13, EI078H06, BI100J15, BI107E19, BI131F09, BI075E21, BI07"		
	/tissue_type="whole plant" /dev_stage="whole-life-cycle" /lab_host="E. coli DH10B" /clone_lib="A normalized whole-life-cycle cDNA library of rice" /notes="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; The library is constructed based on the strategy of saturation hybridization with genomic DNA using rice cultivar Minghui 63. This library consists of cDNA from 15 directionally cloned cDNA libraries constructed with different tissues from 9 developmental stages."		
ORIGIN			
Alignment Scores:			
Pred. No.:	1_99e-93	Length:	1749
Score:	909.00	Matches:	179
Percent Similarity:	62.3%	Conservative:	69
Best Local Similarity:	45.0%	Mismatches:	148
Query Match:	44.4%	Indels:	2

DB:	8	Gaps:	2
US-10-673-786A-2 (1-396) x CX109239 (1-1749)			
Qy	1	MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	281	GTCTTCGCTCGCTCGCGAGCGCGGAGGACCCCATCTCTCGGGGTGACGGTCCGCTAC	340
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr	40
Db	341	AACAAGGATCCAGGCCCGTGAAGTCAATCTCGCGTCCGCGCTACCGGACCGAGAA	400
Qy	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLysLeuGluAsnGlu	60
Db	401	GGGAAGCCCTGGTCTGAATGTGTGAGCGCGCTGAGCAGATCTGATCAACACCCG	460
Qy	61	Thr---ThrLysAsnTyLysGlyIleAspGlyIleProGluPheGlyArgCysThrGln	79
Db	461	TCACGTGTTAAGGATATTTCGCCGATCACTGACTGGCGGATTTCAATAAGCTGAGTCT	520
Qy	80	GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln	99
Db	521	AAGCTTATTTTGGTGTGCTGACAGTCTGCCATTCAAGAGAATAGGTGGCTACAGTTTCAG	580
Qy	100	ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer	119
Db	581	TGCTTGTGAGGAACCTGTTCTTTAAGGTGGAGGTGAATTTCTTGCAGGCACTTATCAT	640
Qy	120	ValLysArgValTrrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer	139
Db	641	GNACGACTATATACATCCACAGCAACCTGGGGGAATCACCCCAAGTGTTCACTTTA	700
Qy	140	AlaGlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAsnHisThrLeuAspPhe	159
Db	701	GCTGGCTGACTTTAGGAGTTACCGCTACTATGATCTCGCAACCCGTGGACTGGATTTC	760
Qy	160	AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly	179
Db	761	CAAGGCTGTAGAAAGATCTCGTTCAGTCTCCTTCAGGTGCAATTTGTACTGCTTCATGCT	820
Qy	180	CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGln	199
Db	821	TGTGCCCAACACCTTACTGGAGTAGACCAACTTTGGACCATGGGAGGAAACACATCAGGCA	880
Qy	200	LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArg	219
Db	881	TTGATGAGATCAAAAGCATTCGTGCATCTTTGATAGCGTTATCAGGGATTTCGAAGT	940
Qy	220	Gly---LeuGluGluAspAlaGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu	238
Db	941	GGAAGTCTTGACCAAGATGCTCAATCAGTCAGCGCATGTTTGTGTGATGGTGGTGAATTG	1000
Qy	239	IleValAlaSerSerTySerLysAsnPheGlyLeuTyTrpAsnGluArgValGlyAlaCys	258
Db	1001	CTCATGGCTCAGAGTACGTACGTAAGAACATGGATGTATGGAGAGCGTGGTGCCTTTA	1060
Qy	259	ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla	278
Db	1061	AGCATAGTTTGGGAAGTGCTGATGATGCTGCTCAGGGTTGAAGTCAACTTAAGCTTGTA	1120
Qy	279	IleArgAlaAsnTySerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeu	298
Db	1121	ATTAGGCTATGATATTCAAAACCTCTTATCATGTGTCATCTATCGTGGCTACCATACTT	1180
Qy	299	SerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle	318
Db	1181	AAGGACATGTCGAATTTCAATGAATGGACTGTGGAGCTGAAGGGCATCGCTGATAGGATT	1240
Qy	319	GlnArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPhe	338
Db	1241	ATTAGCATGAGGCAACACACTTTTGTGCTCTGAAACGAGAGAACTCCTGGAGACTGG	1300
Qy	339	SerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal	358

Db

1301

AGTCACATCATTAAGCAGATCGGAATGTTCACTTTCCTGCGCTCAACAGTGAATG

1360

Qy

359

LeuArgLeuArgGluGluPheGlyValTyAlaAlaSerGlyArgValAsnValala

378

Db

1361

GCITTCATGAGCAGGAATACCATTTACATGACATCTGATGGAGGATCAGTATGGT

1420

Qy

379

GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu

396

Db

1421

GGTTTGAGCGGAGGACCATTCGCGCATCTTTCAGATGCAATCCATGCCGCGATC

1474

RESULT 7

CNS0A2EC

1599 bp

mRNA

linear

HTC 06-FEB-2004

LOCUS

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GS1TPGH60ZA09 of Hormone Treated Callus of strain col-0 of

DEFINITION

Arabidopsis thaliana (thale cress).

ACCESSION

BX828177

VERSION

BX828177.1

GI:42459989

KEYWORDS

HTC; GSLT cDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 1599)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

AUTHORS

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

TITLE

Annotation
Unpublished
2 (bases 1 to 1599)
Genoscope.
Direct Submission

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

REFERENCE

- Web : www.genoscope.cns.fr)

AUTHORS

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.

TITLE

Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.

JOURNAL

URG INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

Location/Qualifiers
1..1599
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTPGH60ZA09"
/tissue_type="Hormone Treated Callus"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
complement(1..1599)
/gene="At4g31990"

gene

Alignment Scores:
Pred. No.: 9.54e-89 Length: 1599
Score: 868.00 Matches: 176
Percent Similarity: 62.2% Conservative: 69
Best Local Similarity: 44.7% Mismatches: 147
Query Match: 42.4% Indels: 2
DB: 4 Gaps: 2

ORIGIN

US-10-673-786a-2 (1-396) x CNS0A2EC (1-1599)

```
QY 2 PheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 218 TTCGAGGGTATAACTATGCTCCACCAGACCCCTATTCTTGGAGTCAGTTAAGCAATTCAA 277
QY 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
Db 278 GCTGACACTAACCGGATGAACCTCAATCTTGGTGTGGTCTTATCGTACTGAGGAATC 337
QY 42 LysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
Db 338 CAGCCTTATGCTTAAATGTTTAAAGGGCGGAGATTTGATGTGTGAGAGAGAGAT 397
QY 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 398 AACAAAGAGTATCTTCCAATTGAGGGTGGCGAGCATTCACAAAGGCTACTGCTGAGTTG 457
QY 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 458 CTATTGGAGCTGGTCATCTCTTATTAGGAACAAGAGTAGCAACAATTCAGGGTCTT 517
QY 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 518 TCGGGAACAGGTCTACTGGATAGCAGCGGTCTTATAGAGCGTATTTCCTCGAGCA 577
QY 122 ArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 578 AAAGTTGTGATCTCATCACCACCTGGGTATATACAGAAATATCTTCAATGATGCCAA 637
QY 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 638 GTTCGCTGGTCCGATACCGCTACTATGATCCAAAACAATGGTTTGGATTTGAGGGA 697
QY 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 698 ATGATAGCAGATATAAGGAAGCTCCAGAAAGATCTTCATCTGCTTCACCGATGTCT 757
QY 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
Db 758 CACACCCCACTGGATTGACCCACACAGACAGTGGTAAATTTGCTGATGTCATT 817
QY 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 818 CAGGAAAAGAACCATATCCATTTTTCGATGTTGTCATACAGGGCTTTCGTAGTGAAGC 877
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 878 CTTGATGAAGATGCGACATCTGTGAGATTATTGCTGAGCGGGGAATGGAGTTTTTGT 937
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluAtqValGlyValaCysThrLeu 260
Db 938 GCTCAGTCATATAGTAAATTTAGTTTGTATGAGAAAGAAATTTGGGCAATCAATGTC 997
QY 261 ValAlaAlaAspSerGlyThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 998 GTGTGCTCTCAGCTGATGCTCTACAGGGTCAAGAGCCAGTTCGAAGAGGATCTCGG 1057
QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1058 CCTATGTACTCGAATCCACAGTTCATGGGGCGGAGAAATCGTGGCCAAATGTAGTGGTGT 1117
QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1118 GTAACTATGTTCCAGTGAATGGAAGCAGAGATGGAAATGATGGCAGGAGAAATAAGACG 1177
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGlnLysGly---AlaAsnArgAspPheSer 339
Db 1178 GTTACACAGAGCTGTATGATAGCTCGTTTCAAAGACAGAGCGGGAGGAGCTGGTCC 1237
QY 340 PheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1238 TTCAATCTGAAGCAAAATGGCATGTTCTTTTCCACGGGCTTAAACAAAGCTCAGAGCGAT 1297
```

```
QY 360 ArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValaAsnValaAlaGly 379
Db 1298 AACATGACGGCAAAATGCGCATGTATATGACTAAAGACGGGAGGATATCATTCGCCGA 1357
QY 380 MetThrProAspAsnMetAlaProLeuCysGluAlaIleVal 393
Db 1358 TTATCTCTGGCCAAATGCGAGTATCTTCTGCTGATCGATCATC 1399

RESULT 8
LOCUS AY105250 1832 bp mRNA linear HTC 18-FEB-2005
DEFINITION Zea mays PCO107699 mRNA sequence.
ACCESSION AY105250
VERSION AY105250.1 GI:21208328
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1832)
AUTHORS Gardiner,J., Schroeder,S., Polacco,M.L., Sanchez-Villeda,H.,
Fang,Z., Morgante,M., Landewe,T., Fengler,K., Useche,F., and Coe,E.H.
Hanafey,M., Tingey,S., Chou,H., Wing,R., Soderlund,C. and Coe,E.H.
Jr.
TITLE Anchoring 9,371 maize expressed sequence tagged unigenes to the
bacterial artificial chromosome contig map by two-dimensional
overgo hybridization
JOURNAL Plant Physiol. 134 (4), 1317-1326 (2004)
PUBMED 15020742
REFERENCE 2 (bases 1 to 1832)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M., and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 3 (bases 1 to 1832)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1. 1832
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeGDB:636247"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,19e-87 Length: 1832
Score: 857.00 Matches: 174
Percent Similarity: 62.5% Conservative: 73
Best Local Similarity: 44.1% Mismatches: 145
Query Match: 41.9% Indels: 3
DB: 4 Gaps: 2
```



```

Db      194  CTTGTTGACTTCGCTTACCAGGGTTTGGCCGGTGGTCTGGAAGAGATGCTGAAGGACTG 253
Qy      229  ArgAlaPheAlaAlaMethHisLysGluLeuLeuValAlaSerSerTyrSerLysAsnPhe 248
Db      254  CGCGCTTCGCGGTATGATTAAGAGCTGATTTGGCCAGTTCCTACTCTAAAACATTT 313
Qy      249  GlyLeuTyrAsnGluArgValGlyAlaCysThrLeuValAlaAlaAspSerGluThrVal 268
Db      314  GCGCTGTACAAACAGCGTGTGGCGCTTGTACTCTGTTGCTGCCGACAGTGAACCGTT 373
Qy      269  AspArgAlaPheSerGlnMetLysAlaAlaLeuArgAlaAsnTyrSerAsnProProAla 288
Db      374  GATCGCGCATTCAGCCAAATGAAGCGCGGATTCGCGCTAACTACTCTAACCCACGACGA 433
Qy      289  HisGlyAlaSerValValAlaThrIleLeuSerAsnAspAlaLeuArgAlaIleTrpGlu 308
Db      434  CAGCGCGCTTCGTGTTGCCACCATCTCGAGCAAGATGCGTTAGTCGGATTTGGGNA 493
Qy      309  GlnGluLeuThrAspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValAsnThr 328
Db      494  CAAGAGTGACTGATGATGCCAGCGTATTTCAGCGTATGCGTCAGTTGTTGCTCAATACG 553
Qy      329  Leu 329
Db      554  CTG 556

CNS09ZWQ      1486 bp      mRNA      linear      HTC 06-FEB-2004
LOCUS      Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION      GSLTFB192F03 of Flowers and buds of strain col-0 of Arabidopsis
                thaliana (thale cress).
ACCESSION      BX829537
VERSION        BX829537.1 GI:42456349
KEYWORDS       HTC; GSLT cDNA.
SOURCE        Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
                rosids; eurosids II; Brassicaceae; Arabidopsi.
REFERENCE      1 (bases 1 to 1486)
AUTHORS        Castell, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,
                Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
                Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE          Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
                A Combined Approach to Evaluate and Improve Arabidopsis Genome
                Annotation
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 1486)
AUTHORS        Genoscope.
JOURNAL        Direct Submission
TITLE          Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT        The sequences are based on single pass reads.
                Life Technologies (a division of Invitrogen) members carried out
                full-length libraries construction : Temple G.
                Genoscope members carried out sequencing and annotation : Castell
                V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
                Schachter V., Weissenbach J., Salanoubat M.
                URV INRA : Clepet C., Caboche M.
                Annotation is based on the June 2003 version of the Arabidopsis
                genome released by MIPS (Munich Information center for Protein
                Sequences). 5 prime and 3 prime are assembled with Phrap.
                http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full-length
                http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
                Location/Qualifiers
                1..1486
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
FEATURES
source

```

```

/clone="GSLTFB192F03"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..1486
/gene="At5g19550"

Alignment Scores:
Pred. No.:      5,57e-85      Length:      1486
Score:          835.00      Matches:      176
Percent Similarity: 60.7%      Conservative: 68
Best Local Similarity: 43.8%      Mismatches: 152
Query Match:    40.8%      Indels:      6
DB:             4          Gaps:      2

US-10-673-786A-2 (1-396) x CNS09ZWQ (1-1486)

Qy      1  MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db      99  GTCTTCTCTAACGTTGCTCGTCTCCGAAAGATCCTATTCTCGGTGTAACCTGTGCTTAC 158
Qy      21  ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db      159  AACATGATCCTAGTCCTGTTAAAGATCAATTTGGGTGTCGGTCCCTACCGAACTGAGAA 218
Qy      41  GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db      219  GGGAAAGCCTCTTGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 278
Qy      61  Thr---ThrLysAsnTyrLeuGlyLysIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db      279  TCTCGGTCGAAGGAATACATCCCATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 338
Qy      80  GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysLysArgAlaArgThrAlaGln 99
Db      339  AAGCTCATCTTAGGTGCTGACAGTCTCTCGCATTTACAGAAAGTAGAGTTACTACAGTCCAG 398
Qy      100  ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db      399  TGCTTGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 458
Qy      120  ValLysArgValTrpValSerAsnProSerTrpProAsnHis-LysSerValPheAsnSe 139
Db      459  CAAAGTGTCATTACATTCCTCAAAACCAACTTGGGGGAACCATCCCAAGATTTTCAACTT 518
Qy      139  rAlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPh 159
Db      519  GGCAGGCTTGTCTGTGGAGTATTTCCGCTACTATGATCTCTGCAACCCCGTGGTCTTGACTT 578
Qy      159  eAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisG1 179
Db      579  CAAAGGCTTGTCTGAGGATCTTGGCGCTGACCATCTCGAGCTATTGCTTACTTCTATGCG 638
Qy      179  yCysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaG1 199
Db      639  ATGTTACACATCCCACTGCGAGTTGACCCCAACCTCTGAACAGTGGGAACAGATTCGACA 698
Qy      199  nLeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaAr 219
Db      699  GCTGATGAGATCTAAAAGCTTATTACCCTTTTGTGATAGTCATATCAGGGTTTGTGCTAG 758
Qy      219  gGly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLe 238
Db      759  TGGTAGCCTTAAACACAGATGCACATCTGTCCTGACATTTGTTGCTGATGGCGGTGAATG 818
Qy      238  uIleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCy 258
Db      819  CTTGATAGTCAAAGTTACGCCAAAATATCGGACTTTATGGGAGCGCTGTTGGTGCCCT 878
Qy      258  sThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAl 278
Db      879  TAGCATTGCTGCAAGTCAGCAGATGCTGCTAGTAGTAAGGTTGAGAGCCAGGTGAAGCTTGT 938

```

```

Qy 278 aileArlaAenTyrSerAen-ProproAlaHisGlyAlaSerValValAlaThrIleL 298
Db 939 TGTGGGCCCATGATTTCGAGCCGCCCTATTTCATGGAGCATCAATTGTCGCCACCATTC 998
Qy 298 eUserAenAlaLeuArlaIleTTPGluGlnGluLeuThrAspMetArgGlnArgI 318
Db 999 TAAAGAGCATGATGTATACAACTGAGCCATCGAGCTGAAGAAATGCTGACCGTA 1058
Qy 318 leGln--ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAs 337
Db 1059 TAAAGCGGCATCGCCCAACAGTATTGAGAGCTATACAGCTAGAGGTACACCTGGTTA 1118
Qy 337 pPheSerPheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluG 357
Db 1119 CTGGAGTCATATTATCAACAGATTGGGATGTTTACATTACTGGATTGAACAAGGACA 1178
Qy 357 nValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnVa 377
Db 1179 AGTGAATTCATGACCAAGAGTTCACATTATACATTTACATGCTCTGATGGAGAAATAGCAT 1238
Qy 377 lalaGlnMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1239 GCGAGTCTAAGTTCGAAGACAGTGCCTCACCTTGCAGTCTATGCGATGCTGCAGTG 1296

RESULT 11
LOCUS AY815146 1502 bp mRNA linear HTC 23-PEB-2005
DEFINITION Schistosoma japonicum clone SJCHGC06163 unknown mRNA.
ACCESSION AY815146
VERSION AY815146.1 GI:56757418
KEYWORDS HTC.
SOURCE Schistosoma japonicum
ORGANISM Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatoides; Schistosomatidae; Schistosoma.
1 (bases 1 to 1502)
Han, Z.
The full-length cDNA sequences of Schistosoma japonicum genes
Unpublished
REFERENCE 2 (bases 1 to 1502)
Han, Z.
Direct Submission
Submitted (02-NOV-2004) Functional Genomic Department, Chinese
National Human Genome Center at Shanghai, 351 Guo Shoujing Road,
Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P.R.China
Location/Qualifiers
1..1502
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/db_xref="taxon:6182"
/clone="SJCHGC06163"
53..1306
/note="similar to aspartate aminotransferase Complex with
Alpha-Methyl (45.6 kD) (XG861)"
/codon_start=1
/product="unknown"
/protein_id="AAW26878.1"
/db_xref="GI:56757418"
/translation="MPTKSPVTVVTKLSSHNHVKLPPDAILGITAEYNRDTNPQ
KINLGAYRDNDKQFVLPVSKAEALLAKNKEYAPISGPQFCDLSIKUALTE
QSPRIKORCNATVQTSQGTALRIGGAFINEAQKHIMWSPTWQHHIFVHSGLN
VQYRYDSNTCGFDIDCLSLDKIPKGFVLLHACAHNPSTGVDPSFDQWRKIGETM
KRDLIIPFDCAQYGFASGDLNNDKAIKYFTDELNPFTLPTQSFKNMGLYGERVG
AFSLICSSDEAECLSQIKILIRPMSNPPIHGARIAAEMLSNVDLRRKWLIDKSM
ADRIITKQSLKNGLIERGSDHSHITNQIMFCFSGNLNPSQVEKLTNEYSIYLTKD
GRISLAGLSSKNVYLAAHQVTK"

CDs
Alignment Scores: 8.55e-82 Length: 1502
Pred. No.: 807.50 Matches: 161
Score: 57.4% Conservative: 65
Percent Similarity:

```

```

Best Local Similarity: 40.9% Mismatches: 165
Query Match: 39.5% Indels: 3
DB: 4 Gaps: 2
US-10-673-786A-2 (1-396) x AY815146 (1-1502)
Qy 2 PheGluAsnThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 107 TGGAAACCATGTTAAGCTGGGACCTCCAGATGCAATTTTGGGAATTTACTGAGCCTACAAT 166
Qy 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
Db 167 CGTGACACCAATCTCGAAAAATTAATCTAGGAGCTGGAGCCTATAGGATGATAATGGA 226
Qy 42 LysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
Db 227 AAACCTTTTCTTGCCATCTGTCAAAGAGCGGAGTCACCTTTTATTAGCGAAAAATTTG 286
Qy 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 287 AATAAGGAGTATGCTCCAATCAGTGGCAATTCGCCAATTTTTCGCAATTTATCCATTAACTT 346
Qy 82 LeuPheGlyLysGlySerAlaLeuLysAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 347 GCATTACTGAACAGTCTCCACGAATCAAAGATCGTTGCAATGCTACGTTACAGCAATA 406
Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 407 TCTGGCACTGCTCTTACGAATTTGGTGAGCTTTTATTAAATGAATTCGCAGAACAAAAA 466
Qy 122 ArgValTrpValSerAenProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 467 CACATTTGGATGCGCTGCACCTACTTGGGGCAATCATATACCCCAATTTTCGTGCATATGGA 526
Qy 142 LeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 527 TTGAATGTACATCAATATCGTTATTATGATAGTAATACATGTGGATTTGATTTATGATGGA 586
Qy 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 587 TGTTTGGCGATTGTCAAATAATTCCTAAAGGACATTTTGTATTATACATGCGATGTGCA 646
Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSer 201
Db 647 CATAATCCTACAGAGTTGATTCATCATTTGATCAATGGAGGAAAAATTTGGAGAGATTATG 706
Qy 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly--- 220
Db 707 AAATCAGTGATTTAATTCATTTTGTGCTTATGCTTATCAAGGATTCGCTTCAGGTGAC 766
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAla-----AlaMetHisLysGluLeu 238
Db 767 ATTAATAATGATCGAAAGCAATAGCTTATTTTACAGACGAATTTGAATTTCCCTACTTTA 826
Qy 239 IleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCys 258
Db 827 TTCCTTACACAAAAGTTTCGCTAAAAATATGCGGATATATACGGGAAACGCTGTGGTCGCTT 886
Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
Db 887 AGTCTCTTTTGTCTTCATCAGATGAAGCGCAACGTTGTTTATACAAATTTAAATACTTT 946
Qy 279 IleArgAlaAsnTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 947 ATTCCACCAATGATAGTAAATCCACCAATTCACGGTGCACGCTATTGCTGCTGAATTAATG 1006
Qy 299 SerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle 318
Db 1007 TCTAATTTGATTATGCTCGAAAAATGGCTCATTTGATTGAAATCTATGGCTGATCGCTATT 1066
Qy 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPhe 338
Db 1067 ATAACAATGCGCCAAAGTTTAAAGATGGTTTAAATCATCGAAGCAGGATCTCAACATGACTGG 1126

```



```

/ gene = "DKF2p469B0125"
join(103..1116,1116..1394)
/ gene = "DKF2p469B0125"
/ codon_start = 1
/ product = "hypothetical protein"
/ protein_id = "CAH92240.1"
/ db_xref = "GI:55731048"
/ translation = "WALLHSGRALPGIAAAPHGLAAASARASSWTHVEMGPPDPPI
LCVTFARFDTNKGSMKNLGVGAYRDDNGKPYVLPSVRKAQIAAKNLDKEVLTGGL
AEFCASAEALGENSEVLKSGRFVTVOTISGTGALRIGASFLQRFKPSRDVLPKLP
SKWNTHTPFRDAGMQLQGYRYDPKTCGDFDFTGAVEDISKIPQSVLLHACAHNPTG
VDPPEQKTEATVVKKNLFAFFDMAYCGFASGDDKDAWVRHFIISGNNCLQCS
YAKNGLGGERVGAFTWCKDAEKRVESQKILIRPYNPNPLNGARIAAILNTP
DLRKTWLEQVGMADRIIGMRTQLVSNLKGSTHWOHTIDQIGMFCFTGLKPEQVE
RLIKEFSYMTKDGRIISVAGVTSNVGVLAHAHQVTK"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 6,03e-78 Length: 2447
Score: 777.00 Matches: 162
Percent Similarity: 58.3% Conservative: 66
Best Local Similarity: 41.4% Mismatches: 161
Query Match: 38.0% Indels: 3
DB: 4 Gaps: 2

```

US-10-673-786A-2 (1-396) x CR860094 (1-2447)

```

QY 4 AsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArgAlaAsp 23
DB 205 CATGTGGAAATGGGACCTCCAGATCCCATTTGGGAGTCACGAAGCCCTTTAAGAGGGAC 264
QY 24 GluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThrGlyLysThr 43
DB 265 ACCAATAGCAAAAGATGATCTGGAGTGTGGCTACCGGATGATACCGAAAGCCT 324
QY 44 ProValLeuThrSerValLysLysAlaGluInTyLysLeuGluAsnGluThrThrLys 63
DB 325 TAGCTGTGCTAGCTCGCAAGCGAGAGCCAGATGGCCGCAAAATTTTGACACAG 384
QY 64 AsnTyLysGlyIleAspGlyLeuProGluPheGlyArgCysThrGlnGluLeuLeuPhe 83
DB 385 GAATACCTGCCCCATCTGGGGAGCTGGCTGAATTTTTCGAAGGCTATCGCAGAACTAGCCCTG 444
QY 84 GlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrProGlyGly 103
DB 445 GTTGAGAACGCAAGCTCTGAAGAGTGGCCGTTTGTCTGCTGTCGAGACCATTTCTGGA 504
QY 104 ThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLysArg--- 122
DB 505 ACTGGAGCCTTAAGGATCGGACCGAGTTTCTGCAAGAGATTTTAAAGTTCAGCCGAGAT 564
QY 123 ValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGlyLeu 142
DB 565 GTCTTTCTCCCAAAACCATCTCTGGGGAACACACACACCATCTTCAGGGATCTGGCATG 624
QY 143 GluValArgGluTyAlaTyTrpAlaGluAsnHisThrLeuAspPheAspAlaLeu 162
DB 625 CAGCTACAAAGGTATCGGTATTATGACCCCAAGACTTGGCGTTTGTGACTTCACAGGTGCT 684
QY 163 IleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCysHis 182
DB 685 GTGGAGGATATTTCAAAATACACAGCAGAGTGTCTTCTGCTGATCGCTCGCCAC 744
QY 183 AsnProThrGlyIleAspProThrLeuGluGlnInTyLysLeuAlaGlnLeuSerVal 202
DB 745 AATCCACGGGAGTGGACCCCGCTCGGAACAGTGGAGGAATAAGCAACAGTGGTGAAG 804
QY 203 GluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArgGly---Leu 221
DB 805 AAAAGGAATCTCTTTGCGTCTCTTGACATGGCGCTACCAAGGCTTTGCCAGTGGTATGGT 864
QY 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisGlyLeuLeuIleValAla 241

```

```

DB 865 GATAGAGATGCTGGGCTGTGGCCACATTCATCGAACAGGGCATTAATGTTGTCTCTCC 924
QY 242 SerSerTyLysAsnPheGlyLeuTyLysAsnGluArgValGlyAlaCysThrLeuVal 261
DB 925 CAATCATATGCCAAGAACATGGCTTATATGGTAGCGGTAGGAGCCCTTCACTATGCTC 984
QY 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
DB 985 TGCAGATGCGGATGAAGCCAAAGAGGTAGAGTCACAGTTCAGAGATCTTGTATCGCTCC 1044
QY 282 AsnTyLysSerAsnProAlaHisGlyAlaSerValAlaThrIleLeuSerAsnAsp 301
DB 1045 ATGTATTCCACCTCCCTCAATGGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
QY 302 AlaLeuArgAlaIleTyrGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
DB 1105 GATTTCGAAA-ACATGGCTGCNAGAGTGAAGGCGATGGCCGCGCATCATTTGGCATG 1163
QY 322 ArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPheIle 341
DB 1164 CGGACTCAACTGGTCTCCAACTCAAGAGAGGGGTTCACCCACCAACTGGCAACACATC 1223
QY 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlnValLeuArgLeu 361
DB 1224 ACCGACCAAAATGGCATGTTCTGTTTTCACAGGCTTAAAGCCCTGAACAGGTGGAGCGGCTG 1283
QY 362 ArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
DB 1284 ATCAAGGAGTCTCCATCTACATGACAAGGATGGCCGATCTCTGTGCGAGGGTCAACC 1343
QY 382 ProAsnMetAlaProLeuCysGluAlaIle 392
DB 1344 TCCAGCAACGTGGGCTACCTTGCCCATGCCATT 1376
RESULT 15
CR616132 1496 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0CAP003Y015 of Thymus of Homo sapiens
(human).
DEFINITION CR616132
ACCESSION CR616132 GI:50496939
VERSION CR616132.1
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1496)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Raydaway Avenue
REFERENCE 2 (bases 1 to 1496)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
location/Qualifiers
1..1496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="CS0CAP003Y015"
/tissue type="Thymus"
/plasmid="pCMVSPORT_6"

```

```
ORIGIN
Alignment Scores:
Pred. No.: 3,43e-76 Length: 1496
Score: 759.00 Matches: 167
Percent Similarity: 57.1% Conservative: 55
Best Local Similarity: 42.9% Mismatches: 158
Query Match: 37.1% Indels: 9
DB: 4 Gaps: 4

US-10-673-786A-2 (1-396) x CR616132 (1-1496)
Qy 6 ThrAlaAlaProAlaAspPro-IleLeuGlyLeuAlaAspLeuPheArgAlaAspGluAr 25
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 TCCGAGGCCAGCGCTGCTGCTGCTTCAAGCTCACTCCGACTTCAGGAGGATCCGGA 112
Qy 25 gProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGlyLysThrProVa 45
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 CCCCGCAAGGTCAACCTGGGAGTGGGAGATATCGCACGGATGACTGCCATCCCTGGGT 172
Qy 45 lLeuThrSerValLysLysAlaGluGlnTyrLeuLeu---GluAsnGluThrThrLysAs 64
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 TTGCCAGTAGTGAAGAAGTGGAGAGAGATTCCTTAATGCAATAGCCTAAATCAGCA 232
Qy 64 nTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeuLeuPheGl 84
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 GTATCTGCCAATCTCGGCGCTGGCTGAGTTCGGAGCTGTCTTCTCTGCTTGGCCTTGG 292
Qy 84 yLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrProGlyGlyTh 104
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 GGATGACAGCCAGCAGCTCAAGGAGAGCGGGTAGGAGGTGTGCAATCTTTGGGGGGAAC 352
Qy 104 rGlyAlaLeuArgValAlaAlaAspPheLeuAlaLys-----AsnThrSe 119
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 AGTGTCACTTCGATTTGGAGCTGATTTCTTAGCGGTGGTACATGGAACAACAAACA 412
Qy 119 rValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSe 139
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 GAACACACCTGTCTATGTGTCTCTACCAACCTGGGAGAAATCACAATGCTGTGTTTCGC 472
Qy 139 rAlaGlyLeu---GluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAs 158
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 TGCTGGTTTTAAAGACATTCGGTCTCTATCGGTACTGGGATGCAGAGAGAGGATTGGA 532
Qy 158 pPheAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHi 178
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
533 CTCCAGGGCTTCCTGAATGATCTGGAGATGCTCCTGAGTTCCTCCATTGTGTCTCCA 592
Qy 178 sGlyCysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAl 198
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
593 CGCCTGTGCACACCAACCACTGGGATTCGCCAATCTCCGAGCAGTGGAAAGCAGATTGC 652
Qy 198 aGlnLeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAl 218
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
653 TTCTGTATGAAGCACCGGTTTCTGTTCCTCTTTTGACTCAGCCTATCAGGGCTTCGC 712
Qy 218 aArgGly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGl 237
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
713 ATCTGGAACCTGGAGAGAGATGCTGGGCCATTCGCTATTTTGTGTCTGAAGGCTTGA 772
Qy 237 uLeuIleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAl 257
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
773 GTTCTTCTGTGCCAGTCTCTTCCAGAACTTCGGGCTCTACAATGAGAGATCGGAA 832
Qy 257 aCysThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAl 277
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
833 TCTGACTGTGTGGAAAAGAACCTCGAGAGATCTTCGAAGTCTCTTCCAGATGGAGAA 892
Qy 277 aAlaIleArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIl 297
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
893 GATGTGGGATTACTTGGTTCATTCCTCCCGCCCGGAGGAGCAGCAATTTGGCCAGCAC 952
Qy 297 eLeuSerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnAr 317
```

Search completed: March 15, 2006, 16:31:40
Job time : 3904 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 03:10:41 ; Search time 390 Seconds
(without alignments)
1804.909 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENTTAAPADPILGLDLF.....VAGMTPDNAPICEAIVAVL 396

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp

-MODEL=frame+ p2n model -DEV=xlp
-Q=/abs/ABSWER_spool/US10673786/runat_14032006_135613_22673/app_query.fasta.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10673786 @CGN 1.1 512 @runat_14032006_135613_22673
-NCPU=6 -ICPU=3 -NO MAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAD=51 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCUS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1331	3	US-08-985-908-23
2	1931	94.4	1209	3	US-09-489-039A-4902
3	1325	64.8	1830121	3	US-09-557-884-1
4	1325	64.8	1830121	3	US-09-643-990A-1
5	1325	64.8	1830121	3	US-10-158-865-1
6	982	48.0	1242	3	US-09-252-991A-12428
7	982	48.0	1329	3	US-09-252-991A-12162
8	982	48.0	24417	2	US-08-846-762-1
9	955	46.7	579	3	US-09-489-039A-4954

c	10	909.5	44.5	1509	3	US-09-252-991A-12968	Sequence 12968, A
	11	909.5	44.5	1632	3	US-09-252-991A-12507	Sequence 12507, A
	12	908	44.4	702	3	US-09-543-681A-3805	Sequence 3805, Ap
	13	872.5	42.7	100848	3	US-09-596-002-39	Sequence 39, Appl
	14	871.5	42.6	1209	3	US-09-540-236-315	Sequence 315, App
	15	844.5	41.3	1338	3	US-09-489-039A-1301	Sequence 1301, Ap
	16	843.5	41.2	1260	3	US-09-543-681A-3123	Sequence 3123, Ap
	17	830.5	40.6	1284	3	US-09-328-352-769	Sequence 769, App
c	18	824.5	40.3	1185	3	US-09-489-039A-1274	Sequence 1274, Ap
	19	794	38.8	2440	3	US-09-976-594-429	Sequence 429, App
	20	759	37.1	1941	3	US-09-220-132-65	Sequence 65, Appl
	21	759	37.1	1941	3	US-09-814-915A-35	Sequence 35, Appl
	22	759	37.1	1942	3	US-09-949-016-1216	Sequence 1216, Ap
	23	759	37.1	2056	3	US-09-566-921-56	Sequence 56, Appl
c	24	726.5	35.5	1230025	3	US-09-198-452A-1	Sequence 1, Appli
c	25	726.5	35.5	1230230	3	US-09-438-185A-1	Sequence 1, Appli
	26	663	32.4	1127	3	US-09-533-559-59	Sequence 59, Appl
	27	546	26.7	912	3	US-09-248-796A-3374	Sequence 3374, Ap
	28	514	25.1	988	3	US-09-533-559-145	Sequence 145, App
	29	510	24.9	426	3	US-09-543-681A-3717	Sequence 3717, Ap
	30	494.5	24.2	900	3	US-09-533-559-4412	Sequence 4412, Ap
	31	486.5	23.8	867	3	US-09-248-796A-3375	Sequence 3375, Ap
	32	365	17.8	423	3	US-09-252-991A-12380	Sequence 12380, A
	33	359.5	17.6	1517	3	US-09-801-874-1	Sequence 1, Appli
c	34	343	16.8	668	3	US-09-533-559-4436	Sequence 4436, Ap
	35	329	16.1	207	3	US-09-489-039A-4873	Sequence 4873, Ap
	36	325.5	15.9	646	3	US-09-328-111-658	Sequence 658, App
	37	322	15.7	423	3	US-09-252-991A-12333	Sequence 12333, A
	38	321.5	15.7	726	3	US-09-248-796A-3377	Sequence 3377, Ap
	39	317.5	15.5	658	3	US-09-533-559-4632	Sequence 4632, Ap
	40	313	15.3	524032	3	US-09-949-016-16928	Sequence 16928, A
	41	313	15.3	524032	3	US-09-949-016-16929	Sequence 16929, A
	42	313	15.3	524032	3	US-09-949-016-16930	Sequence 16930, A
	43	313	15.3	524032	3	US-09-949-016-16931	Sequence 16931, A
	44	313	15.3	529885	3	US-09-949-016-14340	Sequence 14340, A
	45	313	15.3	529885	3	US-09-949-016-14341	Sequence 14341, A

ALIGNMENTS

RESULT 1

US-08-985-908-23
; Sequence 23, Application US/08985908
; Patent No. 6004773

; GENERAL INFORMATION:

; APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI NA
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,908
FILING DATE: 05-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-325659
FILING DATE: 05-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1331 bases
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
STRAIN: JM109
FEATURE:
NAME/KEY: CDS
LOCATION: 10...1197
US-08-985-908-23

Alignment Scores:

Pred. No.: 5,31e-240 Length: 1331
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-08-985-908-23 (1-1331)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 10 ATGTTTGAGAACATTACCGCGCTCTGCCGACCCGATTCTGGCGCTGGCGGATCTGTT 69
Qy 21 ArgAlaAaspGluArgProGlyValIleAenLeuGlyIleGlyValTyrLysAaspGluThr 40
Db 70 CGTCCGATGAACGTCGCCGGAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 129
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 130 GGCAAAACCCCGGTACTACCGAGCGTGAAAGGCTGAACAGTATCTGCTCGAANAATGAA 189
Qy 61 ThrThrLysAenTyrLeuGlyIleAaspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 190 ACCACCAAAATTAACCTCGCATTCAGCGCATCCCTGGAATTTGGTCGTCACACAGAA 249
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaArgThrAlaGlnThr 100
Db 250 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGCACACT 309
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAenThrSerVal 120
Db 310 CCGGGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTCGCAAAAATACACAGGTT 369
Qy 121 LysArgValTrrPvalSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 370 AAGCGTGTGTGGTGAGCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 429
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAlaGluAenHisThrLeuAaspPheAasp 160
Db 430 GGTCTGGAAGTTCGTGAATACGCTATTATGATCGGAAAATCACACTCTTTGACTTCGAT 489
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 490 GCACGTATTACAGCTGAAGTGAAGCTCAGGCTGGCGACGAGTGTCTTCATGGCTGC 549
Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTrrPginThrLeuAlaGlnLeu 200
Db 550 TGCCATAACCAACCGGTATCGACCTACGCTGGCAACCAATGCAAAACACTGGCACACTC 609
Qy 201 SerValGluLysGlyTrrPleuProLeuPheAaspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 610 TCCGTTGAGAAGGCTGTTTACCGCTGTTTGACTTCGCTTACCGAGGTTTTCGCCGCTGT 669
Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
Db 670 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTCGCGGCTATGCATTAAGAGCTGATGTT 729
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260

Db 730 GCCAGTTCTACTCTAAAAAATTGGGCTGTACAAACGAGCGTGTGGCGCTTGACTCTG 789
Qy 261 ValAlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 790 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGC 849
Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 850 GCTAACTACTTAACCCACAGCACAGCGGCTTCTCTTCAGTGGCTGACAAATGAGCAAC 909
Qy 301 AspAlaLeuArgAlaIleTrrPgluGlnGluLeuThrAaspMetArgGlnArgIleGlnArg 320
Db 910 GATGCGTTACGTGCGATTGGGAACAAGAGCTGACTATATGCCCGCAGCGTATTCAGCGT 969
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenArgAaspPheSerPhe 340
Db 970 ATGCGTCAGTTGTTGTCATACGTCAGGAAAAGGCGCAACCCGCGACTTCAGCTTT 1029
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1030 ATCATCAACAGAACCGCATGTTCTCTTCAGTGGCTGACAAAAGAACAGTCTGCGT 1089
Qy 361 LeuArgGluLupheGlyValTyrAlaValAalaserGlyArgValAenValAlaGlyMet 380
Db 1090 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCGTCGCGTAAATGTGGCGGATG 1149
Qy 381 ThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1150 ACACAGATAACATGGCTTCGCTGTGCAAGCGATTGTGGCAGTGCTG 1197

RESULT 2

US-09-489-039A-4902
Sequence 4902, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 4902
LENGTH: 1209
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4902

Alignment Scores:

Pred. No.: 4,07e-226 Length: 1209
Score: 1931.00 Matches: 370
Percent Similarity: 96.5% Conservatives: 12
Best Local Similarity: 93.4% Mismatches: 14
Query Match: 94.4% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-489-039A-4902 (1-1209)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 19 ATGTTTGAAGAACATTACCGCGCGCCAGCCGACCCGATTTTAGGTCTGGCGGATCTGTT 78
Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAaspGluThr 40
Db 79 CGTCCGATGACCGCCCTGAAAAAATTAACCTCGGAATTTGGTGTTCAGAGATGAAC 138
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 139 GGTAAACGCCCGCTTCAGCACCGCTCAAAAAGCAGACGATATCTGCTGAAAAATGAA 198

61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 199 ACAGCTAAAACATATCTGGCATCGATGTTATCTCTGAATTTGGTCTGCACCCAGGAG 258
 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 259 CTGCTGTTCCGTAAAGCAATGCGATTATCGCTGATAAAGCGCCGCGCCGCGAGACA 318
 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 319 CCGCGCGTACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 378
 121 LysArgValTrpValSerLysProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 379 AAACGTGTGTGGTAAGTAAATCCAGCTGCGGCAACCATATAAAGCGTATTCACCTCTGC 438
 141 GlyLeuGluValArgGlyTrpAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 439 GGGCTGGAAGTGGCGAATACGATACGATACGATACGATACGATACGATACGATACGAT 498
 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 499 GGTCTGCTGCCAGCTGACCAAGCCAGCGGCGGCGAGCTGGTACTGTTCCACGGCTGC 558
 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 559 TGCCCAACCCAGCGATATCGATCGACGCTCGATCGATCGATCGATCGATCGATCGATCG 618
 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 619 TCGGTGGAAAGAGCTGCTACCGCTGTTGATTTGCGCTACGAGCGCTTCGCCCGCGGT 678
 221 LeuGluGluAspAlaGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
 679 CTGGAAGAAGATGCTGAAGGCTACCGCTTTTGTCTCTACATAAAGAGCTGCTGCTC 738
 241 AlaSerSerTyrSerLysAsnPheGlyLeuTrpAsnGluArgValGlyAlaCysThrLeu 260
 739 GCCAGCTCTACTCGAAAAAATTTGGCCTGTACACAGAGCGCTCGCGCGCTGCACTCTG 798
 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 799 GTCCGCGCGGATCAGAGACTGTAGACCGCGCTTTCAGTCAGATGAAGTCGGTGTATCCGC 858
 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 859 GCCAACTACTCGAACCCGCTGCGCATGCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTG 918
 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 919 GATCGCTACGGGCAATCTGGGAGCAGGAACCTGACCGATATCGCGCAGCGCATCCAGCGT 978
 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 979 ATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 1039 ATCACCACGAGAGCGATGCTTCATTCACGCGCGCTGACTAAAGAGAGAGTGTGCGC 1098
 361 LeuArgGluGluPheGlyValTrpAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 1099 CTGCGTGAAGAGTTCGCCATCTATGGGTAGCTTCGCGACGCTATCAACGTGGCGGGATG 1158
 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 1159 AGCGCTGCAATATGGCGCGCTGTGCGAAGCATCTGTCGCGCTACTG 1206

RESULT 3

US-09-557-884-1

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

APPLICANT: Fleischmann et al.
 TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS v6.22
 SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/557,884
 FILING DATE: 25-Apr-2000
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995
 ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks
 REGISTRATION NUMBER: 41,971
 REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8504
 TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double

TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Alignment Scores:

Pred. No.: 7,186-147 Length: 1830121
 Score: 1325.00 Matches: 248
 Percent Similarity: 77.8% Conservative: 60
 Best Local Similarity: 62.6% Mismatches: 88
 Query Match: 64.8% Indels: 0
 DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-557-884-1 (1-1830121)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 DB 1684147 ATGTTTGAACATATCAAGCGCGCACCGCATCCATCTTAGGCTTAGCGAAGCATTT 1684206
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 DB 1684207 AAATCCGAACCTCGGAAATAAATCAATTGGGTATTGGGTTTATAAGATGCGCAA 1684266
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 DB 1684267 GGCACACCCCAATTTATCGACGCGGTAAAGAACGCGAAACACGATTTATTGATAAGGAA 1684326
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 DB 1684327 AAAACCAAGAATTTCTGACTATCGATGCTATTCGGGATTTAATACGAACAAACAAAGCA 1684386
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 DB 1684387 CTCCTTTCGTAAAGATTCTGAAGTCATCAATCTAATCGACGAAGACAGTACAAAGT 1684446
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 DB 1684447 TTAGCGGAACAGGTGCATTACGCAATTTATTAAACGCCAAACTAAAGCA 1684506

```

Qy 121 LysArgValTrpValSerAsnProSerTrpProLeuHisLysSerValPheAsnSerAla 140
Db 1684507 CAAATGTTGGTTCAGCACCACTTGGCCAAACCAATGCGATTTCAATGCTGTC 1684566
Qy 141 GlyLeuGluValArgGlyTyrAlaTyrTyrPheAlaGluAsnHisThrLeuAspPheAsp 160
Db 1684567 GGTATGACCAATTCGGAATATCGTTATATGATGCTGAACCAAGCCCTTGATTTGGGAA 1684626
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 1684627 CACTTATTAGAGATTAAAGCCACGACGAGCGATGTTGCTTTTACACGGTTGT 1684686
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCCGACTGGTATTGACCTACTCCAGAACCAATGCAAGATTAGCCGCACTT 1684746
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 1684747 TCAGCTAAATAATGGTTGGTTGCCACTCTTTCGACTTTGCTTATCAAGGTTTAGCCCAACGGA 1684806
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTAGTGCTTTTCGAGCAACCAACCAAGATTATTAGTG 1684866
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTCATTCTCGAAAACTTTGGTTTATATATGAACGCTGTTGGTGCATTTACCCCTT 1684926
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGGCAGAAAAATGCCAGAAATTCGATCAACCTATTAAACCCCAAGTAAATCAATTTATTCGC 1684986
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1684987 ACCCTATTAATCTTAACCCAGTCTTCACGCGGGGACCGTAGCAACAGTATTAAATGAT 1685046
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1685047 GCTCAACTTCGCCAAGAAATGGGAAATGAATTAATCACTGAATGCGTGAACGCATCAAAAAA 1685106
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 1685107 ATGCGTCACTATTTCGTTTACGTTTATTAAGAAATATGCTGTCAGAACCAAGATTTCAGCTTT 1685166
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1685167 ATCATTGAACAAACGGTATGTTTCTTTCAGTGGATTACAGGGGAACAAGTGGATCGT 1685226
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1685227 TTAATAAATGAATTTGCCATTAGCTGTTCTGTTCTGCTGCTATCAACGTTAGCTGGAATC 1685286
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1685287 ACAGAAGATAATATTTCGTTATCTATGTGAAGTATCGTGAAGTACTT 1685334

```

RESULT 4

```

US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

```

```

; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

```

```

Alignment Scores:
Pred. No.: 7,18e-147 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-643-990A-1 (1-1830121)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1684147 ATGTTTGAACATATCAAGCGCGCCAGCCGATCCAATCTTAGGCTTAGCGAAGCATTT 1684206
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 1684207 AAATCCGAACTCGCAAAATAAAATCAATTTGGTATTGGCGTTTATAAAGATGCGCAA 1684266
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 1684267 GGCACAAACCCCAATTTATGACGCGGTAAAGAGCCGAAACGATTTATTGATAAGGAA 1684326
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 1684327 AAAACCAAGAAATTTATCTGACTATCGATGTTTGGCGATTATAACGACAAACAAAGCA 1684386
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 1684387 CTCCTTTTCGGTAAAGATTCTGAAGTCTCAATCTCAATCGAGCAAGACAGTACAAAGT 1684446
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 1684447 TTAGCGGGAACAGGTGCATTACGCATTTCGCGAGAAATTTATTAAACCCCAACCTAAAGCA 1684506
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 1684507 CAAATGTTGGTTCAGCACCACTTGGCCAAACCAATGCGATTTCAATGCTGTC 1684566

```



```

LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)

```

QY 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 1684567 GGTATGACCATTCGTGAATATCGTTATTATGATGCTGACGCAAGCCCTTCATTGGGAA 1684626
 QY 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 1684627 CACTATTAGAAGATTAAAGCAAGCAAGGCGATGTGTGCTTTTACACGGTTGT 1684686
 QY 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 1684687 TGCCATAATCCGACCTGGTATTGACCTCTCCAGAACAAATGCAAGAAATTAGCCGACATT 1684746
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 1684747 TCAGCTAAATAATGGTTGGTGGCACTCTTTGACTTTTATCAAGTTTAGCCCAACGA 1684806
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuVal 240
 Db 1684807 TTAGATGAAGATGCTTATGTTTACGTGCTTTTGCAGCAAAACCAACAAAGAAATTATTAGTG 1684866
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 1684867 GCGAGTTCAATCTCGAAATAATTTGGTTTATATAATGAACGGTGTGGTGCAATTACCCCTT 1684926
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 1684927 GTGGCAGAAAATGCAAGAAATTCATCACTCACTATTAACCAAGTAAATCAATTATTCG 1684986
 QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 1684987 ACCCTATACCTTAACCCAGCTCTCTACGGCGGGCGACCGTAGCAACAGTATTAAATGAT 1685046
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 1685047 GCTCAACTTCGCAAGAAATGGGAAAATGAATTAACCTGAATCGTGAACGATCAAAAAA 1685106
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyValaAsnArgAspPheSerPhe 340
 Db 1685107 ATCGTCACTTATTCGTCAGTATTAAAGAAATATGTTGTCAGAACAAAGATTTACGTTT 1685166
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1685167 ATCATTTGAACAAAACGGTATGTTTCTTCAGTGGATTAAACAGGGAACAAGTGGATCGT 1685226
 QY 361 LeuArgGluLysGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1685227 TTAATAAATGAATTTGCCATTTACGCTGTTCTGTTCTGTCGTATCAACGTCAGTGAATC 1685286
 QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1685287 ACAGAAGATATATTCGCTATCTATGTGAAGATATCGTGAAGTACTT 1685334

RESULT 6
 US-09-252-991A-12428
 ; Sequence 12428, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196, 136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 12428
 ; LENGTH: 1242
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-12428

Alignment Scores:
 Pred. No.: 5, 81e-110 Length: 1242
 Score: 982.00 Matches: 190
 Percent Similarity: 65.7% Conservative: 70
 Best Local Similarity: 48.0% Mismatches: 136
 Query Match: 48.0% Indels: 0
 DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-991A-12428 (1-1242)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 52 CTGTTTCTCGCTGCAAAATGGCAGCGGTGATCCCATCTTGGCTGAACGAAGCCTTC 111
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 112 AACGCCGATACCCGTCGCGGCAAGATCAACCTGGCGGTGGCGGTGATCTACCAAGAGGAG 171
 QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 172 GGGCCGATTCGCTGTCGTCGCGGTCGAGGTCGCGAGAGAGCCGCGCATCGAGGCCAC 231
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 232 GCGCGCGCGGCTACCTCGCATCGAAGGATCGCGCGCTACGACCAGGCGGTACAGAAA 291
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 292 CTGCTGTTTCGGTAAACAGAGTCCGAGCTGCTGGCGCGCGCGGTGTGTGTACAGCCACG 351
 QY 101 ProGlyGlyThrGlyValaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 352 GTCGCGCGGACCGCGCGCTCAAGCTCGCGCGCGCTTCTCAAGCGCTGCTGCCCGAC 411
 QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 412 GCCACCGTGGCCATCAGCGACCGCGAGCTGGGAAAACACCGCGCGCTGTTTCGAAGCGCC 471
 QY 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 472 GGTCTCCCGGTGCAGAACTACCGCTATTACGACCGCGCGCAATGGCGGTGAACCGCGCC 531
 QY 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 532 GCGCTGCTGGAGACCTGAATCCCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 592 TGCCATATCCGACCGCGGTGATCTCGAGCTGGACGACTGGAACAGCGTGGACGCTG 651
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 652 CTCAGGCAAGGCGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuVal 240
 Db 712 ATCGAGGAAGACCGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValAlaCysThrLeu 260
 Db 772 TCCAGCTGTTCTCCAAATCTTCTGCTCTATGGGAACCGGTGCGCGCCCTCTCGATC 831
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 832 GTAACCGAAAACCGCGGATGAATCGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
 QY 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 892 ACCAACTATTGAAACCCACCGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 24417
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-08-846-762-1

Alignment Scores:

Pred. No.: 6,81e-108 Length: 24417
Score: 982.00 Matches: 190
Percent Similarity: 65.7% Conservativity: 70
Best Local Similarity: 48.0% Mismatches: 136
Query Match: 48.0% Indels: 0
DB: 2 Gaps: 0

US-10-673-786A-2 (1-396) x US-08-846-762-1 (1-24417)

```
QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 23597 CTGTTTCTCGCCGCGAAATGACCGCGGTGATCCCATCTTTGGCTGAACGACCTTC 23538
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 23537 AACGCCGATACCCGTCGCGCAAGATCAACCTGGCGGTGGCGGTGATCTACAAACGAGGAG 23478
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 23477 GGGCGCATTCGCTGTCGCGCGGTGCGAGGTCGCGAGAGCGCGCGCATCGAGGCCAC 23418
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 23417 GCGCGCGCGGTACTCGCGATCGAAGGATCGCGCGCTACGACGAGCGGTACAGAAA 23358
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 23357 CTGCTGTTTCGGTAAACGAGTCCGAGTGTGTCGCGCGCGCGGTGTCGACGACCGCC 23298
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 23297 GTCGCGCGCGCGCGCTCAAGCTCGCGCGCGGTGTCGCGCGCGGTGTCGCGCGCG 23238
QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 23237 GCCACCGTGGCCATCAGCGACCGGAGTGGGAAACACCGCGGCTGTTTGAAGCGCC 23178
QY 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 23177 GGTTCCTCCGCGTGCAGAACTACCGCTATTACGACGCGCGCAATGGCGGTGAACCGCG 23118
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 23117 GGCCTGCTGGAAGACCTGATGCCCTCGCGCGCGCGCTGATCGTGGTCTGCGACCGCTGC 23058
QY 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 23057 TGCCATAATCCGACCGCGCTGATCTCGAGCTGGAGCACTGGAACAGAGTGTGGACGTG 22998
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 22997 CTCAGGCCAAGGGCGACGTGCGGTCTCTCGACATCGCTACCGAGGCTTCGGCAACGCG 22938
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 22937 ATCGAGGAAGACCGCGCGGTGCGCTGTTCCGCCAGTGGCGGCTGAGCTTCTTCGTT 22878
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 22877 TCCAGCTGTTCTCCAAATCTTCTCGCTCTATGCGAAGCGGTGCGCGCGCTCTCGATC 22818
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 22817 GTGACCGAAAGCGCGATGAATCGCGCGGTGCTGTCCAGGTGAAGCGGTGATCCGCG 22758
```

```
QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 22757 ACCAACTATTGAAACCCACCGACCCAGCGCGCGCTCTCTCCGTGCTCAACAGC 22698
QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 22697 CCGGAACATGCGGGGCCCTCTGGAGGAGGAGAACTGGGCGGAGATCGCGACCGCATCGCGAC 22638
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 22637 ATGCGCTGGCGATGTCAGCAACTGGCAGCCGCGCGCGCAAGCGGACTTCAGTTTC 22578
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 22577 GTGCGTGCACCAACGTGGCATGTTCTCTATTCCGCGCTGACCGCGGACCGAGTTCGAGCGC 22518
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 22517 CTGAAGACCGAGTTCGGTATCTACCGCTGACACCGCGCGGTATCTGCGTCCGCGCGCTG 22458
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 22457 AACAAATCGAACCTGGAGACCATCACCAAGGCCATCTGTCAGGTCTCTG 22410
```

RESULT 9

US-09-489-039A-4954/c
; Sequence 4954, Application US/09489039A
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4954
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4954

Alignment Scores:

Pred. No.: 3,45e-107 Length: 579
Score: 955.00 Matches: 179
Percent Similarity: 95.3% Conservativity: 5
Best Local Similarity: 92.7% Mismatches: 9
Query Match: 46.7% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-489-039A-4954 (1-579)

```
QY 14 LeuGlyLeuAlaAspLeuPheArgAlaAspGlyArgProGlyLysIleAsnLeuGlyIle 33
Db 579 TTAGTCTGGCGGATCTGTTTCGTCGCGATGACCGCTGAAAAAATAACCTCGGAATT 520
QY 34 GlyValTyrLysAspGluThrGlyLysThrProValLeuThrSerValLysLysAlaGlu 53
Db 519 GGTGTTTACAAGGATGAACCGGTAAAAACGCGCTTCTGACCGCGTCAAAAAACGACAG 460
QY 54 GlnTyrLeuLeuGluAsnGluThrThrLysAsnTyrLeuGlyIleAspGlyIleProGlu 73
Db 459 CAGTATCTGCTGGAAAAATGAACACGACTAAAAAATATCTGGGCGATCGATGGTATTCCTGAA 400
QY 74 PheGlyArgCysThrGlnGluLeuPheGlyLysGlySerAlaLeuIleAsnAspLys 93
Db 399 TTTGGTCTGCTGACCCAGGAGCTGCTGTTTCGTTAAGCAATGCGATTATCGTGATAAA 340
QY 94 ArgAlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPhe 113
Db 339 CGCGCCCGCACCGCGCAGACACCGCGCGGTACCGGTGCGCTGCGGTGCGCGCAGACTTC 280
```

```
Qy 114 LeuAlaLysAsnThrSerValLysArgValTrpValSerAsnProSerTrpProAsnHis 133
Db 279 CTCGCCAAAACACCGAGCTGAAACGCTGTGTGGTAAGTAATCCGAGCTGGCCGAACCAT 220
Qy 134 LysSerValPheAsnSerAlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGlu 153
Db 219 AAAGCGGTATTCCTCTGCGCGGCTGGAAGTGGCGGAATACGCATACGACGCGGCT 160
Qy 154 AsnHisThrLeuAspPheAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAsp 173
Db 159 AACCAACGCTGGACTTTGATGGTCTGCTGGCCAGCTGAACGAAGCCAGCGGGCGAC 100
Qy 174 ValValLeuPheHisGlyCysCysHisAsnProThrGlyIleAspProThrLeuGluGln 193
Db 99 GTGGTACTGTTCCACGGCTGTGCCACCAACCGCGGTATCGATCCGACGCTCGATCAG 40
Qy 194 TrpGlnThrLeuAlaGlnLeuSerValGluLysGlyTyr 206
Db 39 TGGCAGCAGCTGGCCGAGCTGTGCGTGGAAGAGGCTGG 1
```

RESULT 10

US-09-252-991A-12968/c

; Sequence 12968, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 12968

; LENGTH: 1509

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-12968

Alignment Scores:

Pred. No.:	5,91e-101	Length:	1509
Score:	909.50	Matches:	185
Percent Similarity:	62.7%	Conservative:	62
Best Local Similarity:	47.0%	Mismatches:	146
Query Match:	44.5%	Indels:	1
DB:	3	Gaps:	1

US-10-673-786A-2 (1-396) x US-09-252-991A-12968 (1-1509)

```
Qy 2 PheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 1290 TTCCGCAAGTGGCCGGGTACCGGGCGACCGGATCTCTGGGCTCTCGACGCTACCGC 1231
Qy 22 AlaAspGluArgProGlyIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
Db 1230 AACGATCCGCGGGGACCAAGCTCTGGCGTGGTGTCTACAGGATGCCAGGCG 1171
Qy 42 LysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
Db 1170 CTGACCCCGATCTCGCTCGGTGAACTCGCCGAGCAGCGCTGGTTCGACGAGMAACC 1111
Qy 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 1110 ACCAAGACTAGTCTGGCGGCGCACGGGATCGCTGTTCGCGCGGCTGGCGGAATG 1051
Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 1050 GCGTTCGGCGCGGCTCGCCGCTGTGCTGGAGCAACGCGCGGACCGCCAGACGCGCC 991
```

```
Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 990 GCGGCGACCGCGCTTTCGCGCTGGCGGCGACTTCATCGCCCATTCCTGCCCGCGCGC 931
Qy 122 ArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 930 GGCATCTGGCTGAGCGACCGGACCTGGCCCATCCACGAGACCCCTGTCGCCCGCGCGC 871
Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 870 CTGAAGGTTTCCCACTACCCCTACGTCAGCGCGCGCACCAACCGC---CTGGATGTGAGCG 814
Qy 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 813 ATGCTTCTGCTGCGCTGGAGCGCATTTCCCGAGGAGACGTGGTGTGCTGCATGCTGTGTC 754
Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
Db 753 CACACCCGACCGGTTTCGACCTGAGCCACGACGACTGGCGAGGCTGCTCGAGCTGGTG 694
Qy 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
Db 693 CGTCGCGCGAGCTGCTGCGCTGATCGACTTCGCTTACGAGGCTTCGGCGACGGTCTC 634
Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleValAla 241
Db 633 GAGGAAGACGCTGGCGGCTGACGCTGTTTCGCGCGGAACCTGCGGAGGTGCTGCTCACC 574
Qy 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
Db 573 AGTTCCTGCTCGAAGAACTTCGCGCTGTACCGGACCGCTCGGGGCGCTGATCGTCTGC 514
Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db 513 GCGCAGAACCGCGAAGCTCACCGACCTCGTAGCCAACTGGCTTCCTCGCCCGGAAC 454
Qy 282 AsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db 453 CTCTGTCGACCGCGCGCATGCGCGAGGTGTCGCGGATCTCGCGGATCTCGCGGACAGC 394
Qy 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgGlnGlnArgMet 321
Db 393 GAGTTGAAGGAGCTTTGGCAGGAAGGTTCGAGGCGATCGCTCGCCCATCGCCAGCGCTG 334
Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGlyAlaAsnArgAspPheSerPheIle 341
Db 333 CGCATCGGCTGTTCGAAGCGCTGGCGCGCACGCGCTGGCGGAGCGCTTCGCCCATGTC 274
Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 273 GCGCGCACACGCGGATGTTTCTATACCGGACTGAGCGCGGAGGTCGCTCGGCTG 214
Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 213 CGGACGAGCAGCCCGCTTTACCTGGTGTCCAGCGCGCGGCCCAACGTCGCGGTATAGAC 154
Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIleValAlaVal 395
Db 153 GCGCGCGGCTCGACCGCTTCGCGCAAGCATCGCCAGGTC 112
```

RESULT 11

US-09-252-991A-12507

; Sequence 12507, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12507
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12507

Alignment Scores:
Pred. No.: 6,7e-101 Length: 1632
Score: 909.50 Matches: 185
Percent Similarity: 62.7% Conservative: 62
Best Local Similarity: 47.0% Mismatches: 146
Query Match: 44.5% Indels: 1
DB: 3 Gaps: 1

US-10-673-786A-2 (1-396) x US-09-252-991A-12507 (1-1632)

QY 2 PheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 442 TTCGCCAAGTCCGCGCGTACCGGCGACCGCGATCTCGGCTGTCTCGACGCTACCGC 501

QY 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
Db 502 AACGATCCGCGCGGACAAAGCTGGACCTCGGCTCGGTGTCTACAAGGATGCCAGGCG 561

QY 42 LysThrProValLeuThrSerValLysAlaGluGlnTyrLeuGluAsnGluThr 61
Db 562 CTGACCCCGATCTCGGCTCGGTGAACTCGCGGACGCGCTGTGTGACGAGCAACCC 621

QY 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 622 ACCAGAGCTAGTCTCGCGCGCACCGCGATCGCTTTCGCGCGCGCTGTGCGGAACTG 681

QY 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 682 GCGCTCGCGCGCGCTCGCGCTGTGTGCGGACCAACCGCGCGACCCACCGACGCGCC 741

QY 102 GlyGlyThrClyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 742 GCGCGCACCGCGCGCTTTCGCGCTGCGCGCGCTTCATCGCCCGCTGCGCGCGCGC 801

QY 122 ArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAlaGly 141
Db 802 GGCATCTGCTGAGCGACCGCGCTGCGCGCGCGCGATCCAGACCGCTGTTCGCGCGCGC 861

QY 142 LeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 862 CTGAAGTTTCCACCTACCTACCTAGCTCAGCGCGCGCACACCGC--CTGGATGTGAGGCG 918

QY 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 919 ATGCTTGCTGCGCTGAGCGCGCATTCGCCAGGGAGAGCTGTGTGTGTGCTGCTGCTGC 978

QY 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
Db 979 CACAACCCGACCGGTTTCGACCTGAGCCAGCAGCTGCGCGCGAGGCTGTCTGAGGTTG 1038

QY 202 ValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
Db 1039 CGTCGCGCGAGCTGTGCGCTGATCGATCTCGCTACCGAGGCTTCGCGCGAGGCTCTC 1098

QY 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGlnLeuIleValAla 241
Db 1099 GAGGAAGACCGCTGCGCGCTGAGCTGTGTTCGCGCGGAACTGCGCGAGGCTGTGTGTGCT 1158

QY 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValClyAlaCysThrLeuVal 261
Db 1159 AGTCTCTGCTCGAAGAACCTTCGCGCTGTATCCGCGACCGCGCTGCGCGCTGTGCTGCT 1218

QY 262 AlaAlaAspSerGluThrValAspAlaPheSerGlnMetLysAlaAlaIleArgAla 281

Db 1219 GCGGAGAAGCGCGAGAGCTACCGACCTCGGTAGCCAACTGGCCTTCTCSCCGCAAC 1278

QY 282 AsnTyrSerAsnProAlaHisGlyAlaSerValAlaThrIleLeuSerAsnAsp 301
Db 1279 CTCTGGTCGACCCCGCGCGCATGCGCGAGGTGGTTCGCGCGCATCTCTCGCGACAGC 1338

QY 302 AlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 1339 GAGTTGAAGGAGCTTTGGCAGGAAGGTTCGAAGCATCGCTCGCGCATCCGACGCTG 1398

QY 322 ArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPheIle 341
Db 1399 CGCATCGCGCTGTGAAAGCCCTGCGCGCGCACGCGCTGCGCGAGGCTTTCGCCATGTC 1458

QY 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 1459 GCGCGCAACCGCGGATGTTTCTTATACCGGACTGAGCCCGCAGCAGGTCTCGCTGCTG 1518

QY 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 1519 CGCAGCAGCAGCGCGCTTACCTGTGTCTCAGCGCGCGCAACGTCTCGCGTATAGAC 1578

QY 382 ProAspAsnMetAlaProLeuCysGluAlaIleValAlaVal 395

Db 1579 GCGCGCGCGCTCGACCGCTGCGCGCAAGCCATCGCCCGAGGTC 1620

RESULT 12
US-09-543-681A-3805
; Sequence 3805, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3805
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3805

Alignment Scores:
Pred. No.: 2,65e-101 Length: 702
Score: 908.00 Matches: 165
Percent Similarity: 86.0% Conservative: 25
Best Local Similarity: 74.7% Mismatches: 31
Query Match: 44.4% Indels: 0
DB: 3 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-543-681A-3805 (1-702)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 37 ATGTTTGAGAAATCATTTGCTGCACCGCGATCTTATCTGGGTTTACTGATAGTTTC 96

QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 97 CGTCTGATCTCTCGTGAACCAAAATTAACCTAGGATTTGTTTATAGGATGAACA 156

QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 157 GGTAAACACCTGTACTCACCACCGTTTAAAGAGCAGAAAAATATTTATTAGAAACGAA 216

QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 217 TCCACATAAAATATCTTCTTATAGCGGTATCTCTGAGTTTGGTGGCTGCTACTCAAGCA 276

QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100

Db	277	CTGTGTTGGTGAACCAACCCCTATCATCAAGAAAAACGGTGCACGTACAGCACAAAGCT	336
Qy	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAasnThrSerVal	120
Db	337	CCGGGTGGTACTGGGGCATTCAGTATGCTGGGACTTTATTGCTCAGCAAACTAATGCT	396
Qy	121	LysArgValTrpValSerAasnProSerTrpProAasnHisLysSerValPheAasnSerAla	140
Db	397	AAACGTGTATGGATAAGTAACCCGACTTGGCCAAACCAATAATAATATTTTTCAGACTGCA	456
Qy	141	GlyLeuGluValArgGluTyrAlaTyrTyrAaspAlaGluAasnHisThrLeuAaspPheAasp	160
Db	457	GGCTTAGAGATTGGCCAATATGATTATATGATGCCGAAAGCCACGGTTTAGATTTCGAA	516
Qy	161	AlaLeuLeaAasnSerLeuAasnGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys	180
Db	517	GGTATGCTAGCTAGCTACCTACAAAATGCACAGCAGGTGATGTTGCTGTTCCACGGATGT	576
Qy	181	CysHisAasnProThrGlyLeaAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
Db	577	TGCCATAACCAAGTGGTATTGACCCCAACGTTAGAGCAATGGCGTCAATTGGCTGCCTTA	636
Qy	201	SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAargGly	220
Db	637	TCCGCGACAGAAAGTTGGTTTACCAGTATTTGACTTTTGCTTACCAAGGTTTGTCTAATGCG	696
Qy	221	Leu 221	
Db	697	CTT 699	

RESULT 13

US-09-596-002-39

; Sequence 39, Application US/09596002

; Patent No. 6632636

; GENERAL INFORMATION:

; APPLICANT: Lagace, Robert, E.

; APPLICANT: Patterson, Chandra

; APPLICANT: Berg, Kim, L.

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME

; FILE REFERENCE: PM-0008-4 US

; CURRENT APPLICATION NUMBER: US/09/596,002

; CURRENT FILING DATE: 2000-06-16

; PRIOR APPLICATION NUMBER: 60/140,121

; PRIOR FILING DATE: 1999-06-18

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PERL Program

; SEQ ID NO 39

; LENGTH: 100848

; TYPE: DNA

; ORGANISM: M. catarrhalis

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte template ID No. 6632636 39

; PUBLICATION INFORMATION:

US-09-596-002-39

Alignment Scores:		
Pred. No.:	1.65e-93	Length: 100848
Score:	872.50	Matches: 176
Percent Similarity:	59.7%	Conservative: 61
Best Local Similarity:	44.3%	Mismatches: 159
Query Match:	42.7%	Indels: 1
DB:	3	Gaps: 1

US-10-673-786A-2 (1-396) x US-09-596-002-39 (1-100848)

Qy	1	MetPheGluAasnIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe	20
Db	34625	ATGTTTGAACATGTACCACTTATGACGGCGCCCAATTTTAGGCTTGTATGATAAATAT	34684
Qy	21	ArgAlaAaspGluArgProGly---LysIleAasnLeuGlyIleGlyValTrpLysAaspGlu	39

RESULT 14
US-09-540-236-315
; Sequence 315, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 315
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-315

Alignment Scores:
Pred. No.: 1.85e-96 Length: 1209
Score: 871.50 Matches: 176
Percent Similarity: 59.7% Conservative: 61
Best Local Similarity: 44.3% Mismatches: 159
Query Match: 42.6% Indels: 1
DB: 3 Gaps: 1

US-10-673-786A-2 (1-396) x US-09-540-236-315 (1-1209)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 13 ATGTTTGACATGATCAACCACTTATGCGAGCGAGCCCAATTTTAGGCTTGATGATAAATAT 72
QY 21 ArgAlaAspGluArgProGly---LysIleAsnLeuGlyIleGlyValTyrIleAspGlu 39
Db 73 GCCAAGACCCAGCAGCTGACATTAAGTCAATCTTGGTGGCGGTGACTATACCGAA 132
QY 40 ThrGlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsn 59
Db 133 GATGTCATATGCTGCTGACTTGAATCGCTCAAAACCGCGATCCAAATCAACCAATCCA 192
QY 60 GluThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db 193 CCTGTCACGAGGATATTTGCCAATGGATGTTTGGCAGGTTATGCTGAAGCATGCCAA 252
QY 80 GluLeuLeuPheGlyLysSerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln 99
Db 253 GATTTGTTATTTGGTAAATAATCATCTGCGAGTTTGGAGGGCGGTGTTGCCACGATTGCA 312
QY 100 ThrProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db 313 ACTTTGGCGGTTCTGGTGGTAAAGTTGGTGTGATTTTATCATGAGTGGTTCT 372
QY 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
Db 373 ATNGCAATATGCTATGTCACCACTTGGCAATCATCTGGGTATTTTTCAGGGT 432
QY 140 AlaGlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPhe 159
Db 433 GCAGGTTTGGTGGCAAGATGCTTATATGATCCAAATACTTGGCGTAAATTT 492
QY 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 493 GATGAGCTGTCGAATTTTCAAGACTTTAAATGAATAATGATGATGCTACTACACCA 552
QY 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGln 199
Db 553 TGTTCATTAACCACTGTTGGTGGATTTGACCCGTGACACTGGAATGCTACTATTACA 612
QY 200 LeuSerValGluLysGlyTrpLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
Db 613 ATCGTCAAGATATAAACTTATTCATTTATGGATATTCCTATCAAGTTTTCGCGAT 672

QY 220 GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIle 239
Db 673 GATATGGATGGTATGATCGTATGCCATTCGTGCGGGTGGGATGTGACGGTATTT 732
QY 240 ValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThr 259
Db 733 GTCTCTAATTCATTTTCAAAAACCTTGCTACCTTATGTTGAGCGGTGGCGGTTATCT 792
QY 260 LeuValAlaAlaAspSerGlyThrValAspArgAlaPheSerGlnMetLysAlaAlaIle 279
Db 793 GTGCTGTGCACCAACCAAGAGACAGATGTGGTACTTGGTCAATTAATAATTTACCGTC 852
QY 280 ArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSer 299
Db 853 CGTCGATTTTATCAAGTCCAGCTGCACACGCAACTATGTTGTTGATGCGGTCATGAAT 912
QY 300 AsnAspAlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIleGln 319
Db 913 GATGAAGCATTTATTTGCGTTATGATGAGGATATGAAATGCGTGACCGCATTCGT 972
QY 320 ArgMetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSer 339
Db 973 GATATGCTCAAAAACCTCAAGACATTAAGCACTAACTGCCTGAGCGTATTTAGC 1032
QY 340 PheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1033 TATTTTACCAACACGCTGATGTTTACGCTTTACAGGACTAACCAAGAGCAAGTAT 1092
QY 360 ArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGly 379
Db 1093 CGTTTGGCTGAAGAAATTTGCGGTATATATGTTGAAATGCGGTATGTTGTCAGGT 1152
QY 380 MetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1153 CTGAATAATTTCTAATGATAGATATGTTGCCAGTGTATGCGTGAAGTGTCT 1203

RESULT 15
US-09-489-039A-1301
; Sequence 1301, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1301
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1301

Alignment Scores:
Pred. No.: 4.39e-93 Length: 1338
Score: 844.50 Matches: 167
Percent Similarity: 60.7% Conservative: 74
Best Local Similarity: 42.1% Mismatches: 155
Query Match: 41.3% Indels: 1
DB: 3 Gaps: 1

US-10-673-786A-2 (1-396) x US-09-489-039A-1301 (1-1338)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 145 GTGTTTCAAAAGTTGACGCTACCGCGGACCCCTATCTTCTCTGATGGAACGCTTC 204
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 205 AAAGAAGACCCGGAAGCGACAAAGTCAACCTGAGTATCGGCTGTACTACACGATGAC 264

Search completed: March 15, 2006, 04:22:49
Job time : 3161 secs

```
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeu---LeuGluAsn 59
Db 265 GGCATTATCCCGCAGCTGCAGCGGTGCGGAAGCGGAAGCACGCTGAACGCCGAGCGC 324
Qy 60 GluThrThrLysAsnTyrLeuGlyLysLeuGlyLysProGluPheGlyArgCysThrGln 79
Db 325 CATGCGCGCTCGCTGTATCTCCCAATGGAAGGGTTGAGCGGCTACCGGTCAGCGGATTCGC 384
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGln 99
Db 385 CCATCTGCTGTTGGCGCAGACATACCGCGCTTAAGCAAAATCGCATTCCTCTATTTCAG 444
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db 445 ACGGTGCGCGGTTCGGGCGCGCTGAAGGTTCGGGCGGACTTCCTCAAAACGTTACTTTCCT 504
Qy 120 ValLysArgValTTPValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
Db 505 GAGTCTCATGTCTGGGTTCAGCGATCCGACCTGGGAAACACCATCGCCATTTTGAAGGG 564
Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPhe 159
Db 565 GCTGCTTCGAAGTAAGTACTTACCCCTGGTITGATAAGCCACCAACGCGCGTGGCTTT 624
Qy 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 625 GAGGACCTGCTGGCGACGCTCAAAACGCTCGCGCGCGATATTGTGCTGTCACCCCT 684
Qy 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGln 199
Db 685 TGCTGCCACAACCCGACCGCGCGACCTGACGCCAGCCAGTGGGACCGCGTGGTTGAG 744
Qy 200 LeuSerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
Db 745 GTATTGAAGCGCGTACGCTGATCCCTTCCTCGACATCGCCTACGAGGGCTTTGGCGGG 804
Qy 220 GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIle 239
Db 805 GGGCTGGAAGAGGATGCCTACGCTATTTCGCCCATCGCCAGCGCGGATGCCGATGCTG 864
Qy 240 ValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThr 259
Db 865 GTCAGCAACTCGTTCTCCAAATTTTCTCGCTGTACGGGAGCGCGTCCGGCGTCTGTCC 924
Qy 260 LeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIle 279
Db 925 GTGGTCTGTGAAGACAGCGAAACCGCGGCGCGCTGTGGGCGCAGCTGAAGCCACCCTG 984
Qy 280 ArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSer 299
Db 985 CGCCGTAACCTACTCCAGCCCGCCGAGCTTTGCGCGCAGGTGGTGGCGAGCGTGTGAAC 1044
Qy 300 AsnAspAlaLeuArgAlaIleTyrGlnGlnLeuThrAspMetArgGlnArgIleGln 319
Db 1045 GATGCGGGCTGAAAGCCACGCTGGCAGCGGGAAGTGGACGCCATGCGCGCCACATCCTG 1104
Qy 320 ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSer 339
Db 1105 ACCATGCGCGCGCGCTGCTCGACGCGCTGCAGCAGGTTGCCCGGGCGCAAGTGGAT 1164
Qy 340 PheIleLeuLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1165 TACCTGCTCAGCAGCGCGGATGTTTACGCTATACCGGATTTAGCGCGCGCAGGTGGAT 1224
Qy 360 ArgLeuArgGluGluPheGlyValTyrAlaValaSerGlyArgValAsnValAlaGly 379
Db 1225 CGTCTGCGCGATAGTTTGGCGTCTACCTGATTGCGAGCGACGATGTGCGCGCGGG 1284
Qy 380 MetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1285 TTAAACTCACGTAATGTTTACGACGAGGTACGGAAGCTTTTTCGCGCGGTAATG 1335
```

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:51:10 ; Search time 1031 Seconds
(without alignments)
3176.211 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAPADPILGLDLF.....VAGMTPDNAPLCEAIVAVL 396

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10673786/runat_14032006_135621_22855/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fasap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs05h
-USER=US10673786@CGN_1_1_1026@runat_14032006_135621_22855 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1191	6	US-10-369-493-24472
2	2045	100.0	1191	7	US-10-673-786A-1
3	2045	100.0	14759	9	US-10-893-671-87
4	1654	80.9	1188	6	US-10-369-493-44812
5	1343	65.7	7977	9	US-10-795-159-536
6	1343	65.7	908766	9	US-10-795-159-685
7	1325	64.8	1830121	7	US-10-329-670-1
					Sequence 1, Appli

8	1325	64.8	1830121	8	US-10-158-865-1	Sequence 1, Appli
9	1325	64.8	1830121	9	US-10-981-687-1	Sequence 1, Appli
10	1318	64.4	966	6	US-10-369-493-23918	Sequence 23918, A
11	1311	64.1	1191	7	US-10-275-026A-187	Sequence 187, App
12	1310	64.1	17381	9	US-10-915-740A-40	Sequence 40, Appl
13	1310	64.1	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
14	1005.5	49.2	1185	6	US-10-369-493-35511	Sequence 35511, A
15	996	48.7	1188	6	US-10-369-493-31184	Sequence 31184, A
16	996	48.7	1194	6	US-10-369-493-28425	Sequence 28425, A
17	982	48.0	1191	6	US-10-369-493-32111	Sequence 32111, A
18	982	48.0	24417	6	US-10-216-209-1	Sequence 1, Appli
19	975	47.7	1191	6	US-10-369-493-39515	Sequence 39515, A
20	975	47.7	1191	6	US-10-369-493-39895	Sequence 39895, A
21	975	47.7	1194	6	US-10-369-493-39147	Sequence 39147, A
22	947	46.3	1194	6	US-10-369-493-37712	Sequence 37712, A
23	924	45.2	1185	6	US-10-369-493-33145	Sequence 33145, A
24	917	44.8	1185	6	US-10-369-493-32909	Sequence 32909, A
25	917	44.8	1481	7	US-10-425-114-35115	Sequence 35115, A
26	914	44.7	1489	7	US-10-425-114-20320	Sequence 20320, A
27	914	44.7	1817	7	US-10-767-701-14736	Sequence 14736, A
28	913	44.6	1456	7	US-10-425-114-32011	Sequence 22011, A
29	911.5	44.6	1185	6	US-10-369-493-37410	Sequence 37410, A
30	911	44.5	1508	7	US-10-425-114-31281	Sequence 31281, A
31	909	44.4	1203	6	US-10-369-493-41174	Sequence 41174, A
32	909	44.4	2731748	7	US-10-297-465A-1	Sequence 1, Appli
33	904.5	44.2	2152	8	US-10-425-115-12630	Sequence 12630, A
34	893	43.7	1188	6	US-10-369-493-32346	Sequence 32346, A
35	885	43.3	1775	7	US-10-424-599-96239	Sequence 96239, A
36	884.5	43.3	1188	6	US-10-369-493-44841	Sequence 44841, A
37	875.5	42.8	1188	6	US-10-369-493-23974	Sequence 23974, A
38	872.5	42.7	100848	7	US-10-672-787-39	Sequence 39, Appl
39	870	42.5	5976	7	US-10-437-963-84918	Sequence 84918, A
40	866	42.3	1288	8	US-10-425-115-103242	Sequence 103242, A
41	865	42.3	3913	8	US-10-425-115-179006	Sequence 179006, A
42	863	42.2	1932	7	US-10-424-599-29666	Sequence 29666, A
43	861	42.1	1817	7	US-10-767-701-14745	Sequence 14745, A
44	859	42.0	1690	7	US-10-425-114-33582	Sequence 33582, A
45	857	41.9	2084	8	US-10-425-115-179004	Sequence 179004, A

ALIGNMENTS

RESULT 1

US-10-369-493-24472
; Sequence 24472, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24472

Alignment Scores:
Pred. No.: 6.32e-227 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-24472 (1-1191)

QY 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 1 ATGTTTGAAGACATTACCGCGCTCTCCGACCCGATTTCTGGCGCTGGCGGATCTGTTT 60
 QY 21 ArgAlaAaspGluArgProGlyLyseAenLeuGlyIleGlyValTyrlYsAaspGluThr 40
 Db 61 CGTCCGATGAACCTCCCGCAAAATTAACCTCGGGATTTGTTCTATAAAGATGAGACG 120
 QY 41 GlyLyseThrProValLeuThrSerValLyseAlaGluGlnTyrlYsLeuGluAenGlu 60
 Db 121 GGCAAAACCCCGCTACTGACCGCTGCAAAAGCTGAAACAGTATCTGCTATTAAGATGAGACG 180
 QY 61 ThrThrLyseAenTyrlYsLeuGlyIleAaspGlyIleProGluPheGlyValTyrlYsGlnGlu 80
 Db 181 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTCGAAATTTGGTCTGCTGCACTCAGGAA 240
 QY 81 LeuLeuPheGlyLyseGlySerAlaLeuIleAenAspLyseArgAlaAargThrAlaGlnThr 100
 Db 241 CTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGCTGCTGCACGGCACAGACT 300
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaAaspPheLeuAlaIlyseAenThrSerVal 120
 Db 301 CCGGGGGGACCTGGCGCACTACGCGGTGCGGATTTCTTCGCAAAAATACACGCTT 360

US-10-673-786A-1

QY 361 LeuArgGluGluPheGlyValTyrlYsAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCAAGACGTTTGGCGTATATGCGTTGCTTCTGTCGCGTAAATGTCGCGGATG 1140
 QY 381 ThrProAspAsnMetAlaProLeuCyseGluAlaIleValAlaValLeu 396
 Db 1141 ACACAGATAACATGCTCCGCTGCGAAGCGATTTGTGGCAGTGTCTG 1188

RESULT 2

US-10-673-786A-1
 ; Sequence 1, Application US/10673786A
 ; Publication No. US20040132165A1
 ; GENERAL INFORMATION:
 ; APPLICANT: AKHVERDIAN, VALERY ZAVENOVICH
 ; APPLICANT: SAVRASOVA, EKATERINA ALEKSEEVNA
 ; APPLICANT: KAPLAN, ALLA MARKOVNA
 ; APPLICANT: LOBANOV, ANDREY OLEGOVICH
 ; APPLICANT: KOZLOV, YURI IVANOVICH
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-THREONINE USING BACTERIA
 ; TITLE OF INVENTION: BELONGING TO THE GENUS ESCHERICHIA
 ; FILE REFERENCE: US-115
 ; CURRENT APPLICATION NUMBER: US/10/673,786A
 ; CURRENT FILING DATE: 2003-09-30
 ; PRIOR APPLICATION NUMBER: PCT/JP03/02067
 ; PRIOR FILING DATE: 2003-02-27
 ; PRIOR APPLICATION NUMBER: 2002104983
 ; PRIOR FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1191
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1191)
 ; US-10-673-786A-1

Alignment Scores:
 Pred. No.: 6,32e-227 Length: 1191
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservativity: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: Gaps: 0

US-10-673-786A-2 (1-396) x US-10-673-786A-1 (1-1191)

QY 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 1 ATGTTTGAAGACATTACCGCGCTCTCCGACCCGATTTCTGGCGCTGGCGGATCTGTTT 60
 QY 21 ArgAlaAaspGluArgProGlyLyseAenLeuGlyIleGlyValTyrlYsAaspGluThr 40
 Db 61 CGTCCGATGAACCTCCCGCAAAATTAACCTCGGGATTTGTTCTATAAAGATGAGACG 120
 QY 41 GlyLyseThrProValLeuThrSerValLyseAlaGluGlnTyrlYsLeuGluAenGlu 60
 Db 121 GGCAAAACCCCGCTACTGACCGCTGCAAAAGCTGAAACAGTATCTGCTATTAAGATGAGACG 180
 QY 61 ThrThrLyseAenTyrlYsLeuGlyIleAaspGlyIleProGluPheGlyValTyrlYsGlnGlu 80
 Db 181 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTCGAAATTTGGTCTGCTGCACTCAGGAA 240
 QY 81 LeuLeuPheGlyLyseGlySerAlaLeuIleAenAspLyseArgAlaAargThrAlaGlnThr 100
 Db 241 CTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGCTGCTGCACGGCACAGACT 300
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAaAaspPheLeuAlaIlyseAenThrSerVal 120
 Db 301 CCGGGGGGACCTGGCGCACTACGCGGTGCGGATTTCTTCGCAAAAATACACGCTT 360

```
QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAGCGTGTGGTGGAGCAACCAAGCTGGCCGACCAATAGAGCGTCTTTAACTCTGCA 420
QY 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTTCTGTAATAGCTTTATATGATGCGGAAATACACACTCTTGACTTCGAT 480
QY 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCATCGATTACAGCGCTGAATGAAGCTCAGGCTGGCGACGATGCTGTTCCATGGCTGC 540
QY 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCCACACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCACAACTC 600
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlnPheAlaArgGly 220
Db 601 TCCGTTGAGAAGGCTGGTTACCGCTGTTGACTTCGCTTACAGGGTTTTGCCCGTGGT 660
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuVal 240
Db 661 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTCCGCGCTATGCATAAAGAGCTGATTGTT 720
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValaCysThrLeu 260
Db 721 GCCAGTCTCTACTCTAANAACCTTGGCTGTACACAGAGCGTGTGGCGCTGTACTCTG 780
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGTGCGCGACAGTGAACCGTTGATCGCGATTTCAGCAAAATGAAAGCGCGATTGCG 840
QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValAlaThrIleLeuSerAsn 300
Db 841 GCTAACTACTCTAAACCCACAGCACACAGCGCTTCTGTTGTGGCCACCATCTTGAGCAAC 900
QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GATCGTTACGTGGCATTTGGGAACAAGCTGACTGATATGCGCACGCGTATTACAGCGT 960
QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnAspPheSerPhe 340
Db 961 ATCGTCAAGTTCTCGTCAATACGCTGCAGGAAAGGCGCAACCGCGACTTCAGCTTT 1020
QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATCATCAACAGACAGCGCATGTTCTCTTCACTGGCGCTGACAAAAGAACAAAGTGTGCGT 1080
QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 CTGCGCGAAGAGTTTGGCGGTATATGCGGTGTGTTCTGCTGCGGTAAATGTGGCCGGGATG 1140
QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACCAAGATACATGGCTCCGCTGTGGAGGCGATTGTGGCAGTCTG 1188
```

RESULT 3

```
US-10-893-671-87/c
; Sequence 87, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PKZ-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 87
; LENGTH: 14759
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-893-671-87
```

```
Alignment Scores:
Pred. No.: 2,46e-225 Length: 14759
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0
```

US-10-673-786A-2 (1-396) x US-10-893-671-87 (1-14759)

```
QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 2755 ATGTTTGAGAACATTTACCGCGCTCTCCGCCACCGATTCGGGCTCGCCGATCTGTTT 2696
QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 2695 CGTCCCGATGACGTCCTCCGCGCAAAATTAACCTCGGATTTGTTCTATAAAGATGAGACG 2636
QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 2635 GGCAAAACCCCGGTACTGACCAGCGTGAAGAGGCTGAACAGTATCTCTCGAAAATGAA 2576
QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 2575 ACCACCAAAATTAACCTCGGCAATTCACGGCATTCCTGAAATTTGGTCGCTCAGCA 2516
QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 2515 CTGCTGTTTGTAAAGTAGGAGCGCTGATCAATGACAAACGCTGTCGCCACGACAGACT 2456
QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 2455 CCGGGGGGCACCTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATACACAGCGTT 2396
QY 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 2395 AAGCGTGTGGTGGAGCAACCAAGCTGGCGCAACCAATAGAGCGCTCTTTAACTCTGCA 2336
QY 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 2335 GGTCTGGAAGTTCTGTAATACGCTTTATGATGCGGAAATTCACACTCTTGACTTCGAT 2276
QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 2275 GCATCTGATTACAGCTGAATGAAGCTCAGCTGGCGACGTAGTGTCTTCATGGCTGC 2216
QY 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 2215 TGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACCAATGGCAACACTGGCACAACTC 2156
QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 2155 TCCGTTGAGAAGGCTGGTTACCGCTGTTTACCTTCGCTTACCGGGTTTTGCCCGTGGT 2096
QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
Db 2095 CTGGAAGAGATGCTGAAGGACTGGCGCTTTCCGGGCTATGCATAAAGAGCTGATTGTT 2036
QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 2035 GCCAGTCTCTACTCTAAAACCTTTGGCCTGTACACAGCGGTGTGGCGCTTGTTACTCTG 1976
QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1975 GTTGTGCGCGACAGTGAACCGCTTGTATCGCGCATTCAGCCAAATGAAAGCGCGGATTGCG 1916
QY 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
```

Db 1915 GCTAACTACTCTAACCCACAGACACGCGCTTCTGTGTGTGACCATCTCAGCAAC 1856
Qy 301 AspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1855 GATGGTTACGTGGATTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGGT 1796
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLeuGlnGlyAlaAsnArgAspPheSerPhe 340
Db 1795 ATGGCTCAGTTGTTCTGATACGTGTCAGGAAAGGCGCAAAACCGCGACTTCAGCTTT 1736
Qy 341 IleIleLeuGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlnValLeuArg 360
Db 1735 ATCATCAACAGACGCGATGTTCTCTTCAGTGGCTTGACAAAGAACAAAGTCTCGGT 1676
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1675 CTGGCGAAGAGTTGGGTATATGCGGTGCTTCTGTCGCGTAAATGTGGCGGGATG 1616
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1615 ACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGTG 1568

RESULT 4

US-10-369-493-44812
; Sequence 44812, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44812
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-44812

Alignment Scores:
Pred. No.: 1,71e-181 Length: 1188
Score: 1654.00 Matches: 310
Percent Similarity: 86.9% Conservative: 34
Best Local Similarity: 78.3% Mismatches: 52
Query Match: 80.9% Indels: 0
Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-44812 (1-1188)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAAAATTTACAGCAGCACCCTGCCACCCCTATTCTTGTGTAGCGGATAGCTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
Db 61 AAAGCTGATCTCGTGAAATAAATAATCAATCTGGGTATCGGTGTTTACAAAGACGAACT 120
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyLeuLeuGluAsnGlu 60
Db 121 GGTAACCCCTGCTCCTCACCACCGTTAAAGAAAGCAGAAATTCCTGCTCGAAACCGAA 180
Qy 61 ThrThrLysAsnTyLeuGlyLysIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTTATCGGCAATTAGTGGGTACCTGAAATTTGGCGGCTCACTCAGGA 240

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTACTCTTCGGCAATACCAGCATCTATCTACTGACAAACGCGCCGCTACTGTACAAGC 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CAGGCGGAACCGGTGACCTGCGCAGATTTTATTGCTAAACAGCAATGCA 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAACGTCTCTGGATCAGTAACCAACATGCGCAACATAAAGCGGTTTTCTCTAGCGCT 420
Qy 141 GlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGGTTAGAAATCCGCGAGTACCAATATTCAATGCAAAAGCATGCTCTGGATTTGAT 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GGCATGTTGGCTAGCTGCTGAAGCTCAAGCAGGTGATGTTGTTCTGCTCACGGTTGC 540
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCATAAACCAACCGGTATCGATCCAAACGCGTGAACAGTGGCAGAAACCTGGCAGATTTA 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArgGly 220
Db 601 TCTGCGGCAATGGCTGGCTGCTGTTGATTTCCTTATCAGGGTTTTTCCAGAGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 TTAGATGAAGATGCAAGAGCGCTGCGTATTTTACTTAAACCATATGAATGATCGTT 720
Qy 241 AlaSerSerTyrrSerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGCTCTTACTCCAAAAATTTGGCTGTACAAATGAGCGTGTAGTGTGCTGCACCATC 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGCAACTCACAGCGATACCGCAAAAGAGCATTCAGTCAGCGCAATCTATTGTTCTGT 840
Qy 281 AlaAsnTyrrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 ACTAACTACTCTAACCCACCGCACAGCGCTCTCTGTTGTTACTACAATTTCTGTCCAAT 900
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GACGAATTTAAAGCAGAAATGGATTCAAGAACTGCGGACTATGCGTGAACGCAATTCGTCGT 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 961 ATGCGTCAGCTGTTTGTGAACACCTTGCAAGAAAGAGGTGCAAAACAGACCTTTAGCTTT 1020
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATTATTAGCAAAATGATGTTTTCATTTCAGTGTCTGACTTAAAGAAACAAAGTTGATCGT 1080
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 CTGCGTGAAGAATTCGGTATTTATGCGTCAAGTCTGCGGCGCATTAACGTAGCGGGCTTA 1140
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACGTGAGAACATGGTCTCTGTGTGAAGCGGATTTGTCAGTACTC 1188

RESULT 5

US-10-795-159-536
; Sequence 536, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159

```

; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 536
; LENGTH: 7977
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4108)..(4108)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7960)..(7960)
; OTHER INFORMATION: n = a, c, g, or t
US-10-795-159-536

Alignment Scores:
Pred. No.: 3,77e-144 Length: 7977
Score: 1343.00 Matches: 251
Percent Similarity: 78.3% Conservative: 59
Best Local Similarity: 63.4% Mismatches: 86
Query Match: 65.7% Indels: 0
Gaps: 0
DB:

US-10-673-786A-2 (1-396) x US-10-795-159-536 (1-7977)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1594 ATGTTGAACATATTAAAGCGGCACCGATCCCAATCTTAGGCTTAGGCGAAGCATTC 1653
Qy 21 ArgAlaAspGluAtpProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 1654 AAATCCGAACCTCGCGAAATAAATAATTAACCTTGGGTATTGGCGTTTAAAGATCGCGAA 1713
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuGluAsnGlu 60
Db 1714 GGCACACCCCGCATTTATCGTGCATGTAAGAGCAGAAACGCTTATTGATGAAGAA 1773
Qy 61 ThrThrLysAsnTyrLeuGlyLysLeuGlyLysLeuGlyLysLeuGlyLysLeuGly 80
Db 1774 AAAACGAAATATATCTGACTATCGATCGTATGCTGATTAACGAAACAAACAAAGCG 1833
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLysAspLysArgAlaArgThrAlaGlnThr 100
Db 1834 TTACTTTTCGGTAAAGATTCTGAAGTCATCAATCTAATCGTGCAGAACAGTACAAAGT 1893
Qy 101 ProGlyGlyThrGlyValAlaArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 1894 TTAGCGGAACAGCGCCATTCAGCATTCGCGAGAAATTTATTAACGTCACAACTAAAGCA 1953
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 1954 CAAATGTTTGGATCAGCAGCCCACTTGGCCAAACCAATGCCATTTTCAATGCCGTC 2013
Qy 141 GlyLeuGluValArgGlyTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 2014 GGTATGACAATTCGTGAATATCGTTATTATGATGCTGAACGCAAGCCCTTGATTGGAT 2073
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 2074 CATTTACTTGAAGATTTAAGCCCAACAGCAGGAGCGATGTGGCTTTTACCGGTTGT 2133
Qy 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 2134 TGCCATATCCGACTGGTATTGACCTTACTCCAGAACCAATGCCAAGATTAGCCGCACTT 2193
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 2194 TCGGCGAAACAGCGGTGGTGGCACTCTTTCAGCTTTCGCTTATCAAGGTTTAGCCCAATGGA 2253

```

```

Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 2254 TTAGATGAAGATGCTTATGCTGCTTTTGGCGAAACACCAAGAAATTTGTAGTA 2313
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrArgGluValGlyValAlaCysThrLeu 260
Db 2314 GCGAGTTTCATTTCTCGAAACCTTTGGTTTATATGAACGTTTGGTGGCTTTACTCTT 2373
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 2374 GTGCAGAAATGCGAGAAATTCATCAACCGCATTAACCCAGTGAATCAATATTCTGT 2433
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyLysAlaSerValValAlaAlaThrIleLeuSerAsn 300
Db 2434 ACACCTACTCTTAACCTCTCATCTCAGCGCGCGCACTGTAGCAACAGTATTAAATGAC 2493
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 2494 ACTCAACTTCGCCAAGAGTGGGAAATGAATTAACCTGAATCGCGCAACGATCAAAAAA 2553
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAlaAsnArgAspPheSerPhe 340
Db 2554 ATGGTCAATTTATTCGTTCACTTATTAAAGAAATACGCTGGCGAACAAGATTTCACTTT 2613
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 2614 ATCATTTGAACAAACCGGTATGTTCTCTTCACTGATTAACCTGGGGAACAAGTGGATCGC 2673
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 2674 TTAAGAAGAATTTGCCATTTATGCTGCTTCTGCTGCTATCAATGTGGCTGCTATT 2733
Qy 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 2734 ACTGAAGATAATATTCGCTATTATGTGAAGCAATTTGGAAGTGCTT 2781

RESULT 6
US-10-795-159-685/c
; Sequence 685, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 685
; LENGTH: 908766
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9223)..(9223)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (39640)..(39640)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292404)..(292404)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292495)..(292495)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (908604)..(908641)

```

OTHER INFORMATION: n is a, c, g, or t
US-10-795-159-685

Alignment Scores:

Pred. No.: 3,71e-141 Length: 908766
Score: 1343.00 Matches: 251
Percent Similarity: 78.3% Conservative: 59
Best Local Similarity: 63.4% Mismatches: 86
Query Match: 65.7% Indels: 0
DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-795-159-685 (1-908766)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 217833 ATGTTGAACATATTAAGCGGCACAGCCGATCCAATCTTAGGCTTAGGCGAAGCATTC 217774
Qy 21 ArgAlaAaspGluAtpProGlyLysIleAenLeuGlyIleGlyValTyLysAaspGluThr 40
Db 217773 AAATCCGAAACTCCGCGAAATAAATAATTAACCTGGGTATTTGGCGTTTATAAAGATGCCGAA 217714
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAasnGlu 60
Db 217713 GGCACAACCCGATATATCGTCGATGAAAGAGCAAGAAACCGCTTATTGTATAAGGAA 217654
Qy 61 ThrThrLysAenTyLeuGlyIleAaspGlyIleProGluPheGlyA-gCysThrGlnGlu 80
Db 217653 AAAACGAAAAATATCTGACTATCGATGATGCTGATTAACGAAACAAACAAAGCG 217594
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAasnAplysArgAlaAargThrAlaGlnThr 100
Db 217593 TTACTTTTCGTAAGATCTTGAAGTCATCAAAATCTAATCGTGCAGAACAGTCAAAAGT 217534
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAasnThrSerVal 120
Db 217533 TTAGCGGAAACAGCGCATTAAGCATTCGCGCAAGATTTATAACGTCAAACTAAAGCA 217474
Qy 121 LysArgValTrpValSerAasnProSerTrpProAenHisLysSerValPheAasnSerAla 140
Db 217473 CAAAATGTTGGATCAGCACCCCAACTTGGCCAAACCAACCAATGCCATTTTCAATGCCGTC 217414
Qy 141 GlyLeuGluValArgGluTyAlaTyTrpAlaGluAenHisThrLeuAaspPheAasp 160
Db 217413 GGTATGCAATTCGTGAATATCGTTATATGATGCTGAACGCAAAAGCCCTTGATTTGGGAT 217354
Qy 161 AlaLeuIleAenSerLeuAasnGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 217353 CATTTACTTGAAGATTNAGCCAGCAAGCGGATGGTGGCTTTACACGGTTGT 217294
Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 217293 TGCCATAATCCGACTGGTATTGACCCCTACTCCAGAACAAATGGCAAGATTTAGCCGCACTT 217234
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyTrpGlnGlyPheAlaAargGly 220
Db 217233 TCGCGAAAAACGGGTGGTGGCTCTTTGACTTTGCTTATCAAGGTTTAGCCCAATGGA 217174
Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
Db 217173 TTAGATGAAGATGCTTATGGCTTGGCTTTTCGGCGAACCCACCAAGAAATTTAGTA 217114
Qy 241 AlaSerSerTySerLysAasnPheGlyLeuTyTrpAasnGluArgValGlyAlaCysThrLeu 260
Db 217113 GCGAGTTTCATCTCGAAAAACTTTGGTTTATATATGAACGTGTTGGTGGCTTTACTCTT 217054
Qy 261 ValAlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 217053 GTGGCAGAAAAATGCAGAAATTCACACCCGATTAACCCCAAGTGAATCAATTTATTCGT 216994
Qy 281 AlaAenTySerAasnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAasn 300
Db 216993 ACACCTACTACTACCCCTGCATCTCAGCGGCGGCAACTGTAGCAACAGTATTAAATGAC 216934

Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgGlnArg 320
Db 216933 ACTCAACTTCGCCAAGAGTGGGAAATGAATTAACATAATGCGCAACGATCAAAAAA 216874
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAaspPheSerPhe 340
Db 216873 ATGGTCATTTATTCGTTTCAGTTTATAAAGATACCGTGGGACCAAGATTTTCAGCTTT 216814
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 216813 ATCAATTGAACAAAAACGATATGTTCTCTTTTCAGTGGATTAATCGGGGAAACAAGTGCATCGC 216754
Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 216753 TTAAGAAGAATAATTCCTCAATTTATGCTGTTGCTGATCAATGTGGCTGGTATT 216694
Qy 381 ThrProAaspAasnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 216693 ACTGAAGATNATATTCGCTATTATTTATGTGAAGCATTTGTGAAGTGCTT 216646

RESULT 7

US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCES: P8186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature

LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152530)..(152530)

Alignment Scores:
Pred. No.: 1,27e-138 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 7 Gaps: 0

US-10-673-786a-2 (1-396) x US-10-329-670-1 (1-1830121)

Qy 1 MetPheGluAenThrAlaAlaProAlaAspProLeuLeuGluLeuAlaAspLeuPhe 20
Db 1684147 ATGTTGAACATATCAAGCGGCACCGCATCCATCTTAGCTTAGGCGAAGCATTT 1684206
Qy 21 ArgAlaAspGluAargProGlyValIleAsnLeuGluLeuGlyValTyLysAspGluThr 40
Db 1684207 AAATCCGAACCTCGCGAAATTAATCAATTTGGGTATTGGCGTTTATAAAGATGCGCAA 1684266
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 1684267 GGCACAAACCCCAATATATCGCGGTAAAGACCGCGAATATTTGTAAGGAA 1684326
Qy 61 ThrThrLysAsnTyrLeuGlyLysAspGlyLeuProGluPheGlyArgCysThrGlnGlu 80
Db 1684327 AAACCAAGAATATCTGACTATCGATGTTTGGGATTTATAACGAACAAACAAAGCA 1684386
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 1684387 CTCCTTTTCGGTAAGATTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 1684446
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 1684447 TTAGCGGACAGGTGCAATGATGCGGACGAATTTATTAACGCCAAACATAAGCA 1684506
Qy 121 LysArgValTyrPValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 1684507 CAATATGTTGGTGCAGCAGCACCACTTGGCCAAACCAATGCGATTTCAATGCTGTC 1684566
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 1684567 GGTATGACCATTCGTGAATATCGTTATGATGCTGAACGCAAGCCCTTGATTGGGAA 1684626
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValPheHisGlyCys 180
Db 1684627 CACTTATTAGAAGATTTAAGCCCAAGCAAGCGGATGTTGCTTTACACGGTTGT 1684686
Qy 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCCGACTGTTGATGACCTACTCCAGACAAATGCGAAGATTTAGCCGCACTT 1684746
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 1684747 TCAGTAAATATGTTGGTTCGCACTCTTTGACCTTTGCTATCAAGGTTTAGCCACGGA 1684806
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGTTTACGTTTTCGCGCAACCAACCAAGAAATTTATAGTG 1684866
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValAlaCysThrLeu 260
Db 1684867 GCGAGTTCATCTCGAAAAAATTTGGTTTATATATGACGTGTTGGTGCACTTTACCCCTT 1684926
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGGCAGAAAATGCGAAGATTTGCATCAACCTATTAACCAAGTAAATCAATTTATTCGC 1684986
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyLysAlaSerValValAlaThrIleLeuSerAsn 300
Db 1684987 ACCCTATCTACTAACCCAGCTTCTCACGGGGGGCGGCGGTAGCAACAGTATTAAATGAT 1685046

```

```

Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1685047 GCTCAACTTCGCCAAGAAATGGAAAATGAATTAACCTGAAATGCGTAACGATCAAAAA 1685106
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyValAlaAsnArgAspPheSerPhe 340
Db 1685107 ATCGCTCACTTATTCGTTTCAGTTTATTAAGAATATGTTGTCAGAACAAAGATTTTTCAGCTTT 1685166
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1685167 ATCATTGAACAAACGCGTATGTTTCTTTCAGTGGATTACACGGGGAACAAGTGGATCGT 1685226
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1685227 TTAATAAATGAATTTGCCATTTACGTTTTCGTTTCTGTCGTATCAACGTTAGCTGGAATC 1685286
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1685287 ACAGAAGATAATATTCGCTATCTATGTGAAGATATCGTGAAGTACTT 1685334

RESULT 8
US-10-158-865-1
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```

NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

LOCATION: (107248)..(107248)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)

; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Alignment Scores:
Pred. No.: 1,27e-138 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 8 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-158-865-1 (1-1830121)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1684147 ATGTTTGAACATATCAAGCGGACCGCATCCATCTTAGGCTTAGCGAAGCAATTT 1684206
Qy 21 ArgAlaAspGluArgProGlyLyLeAsnLeuGlyIleGlyValTyLysAspGluThr 40
Db 1684207 AAATCCGAAACTCGCGAAATAAATCAATTTGGTATTGGCGTTTATAAAGATGCGCAA 1684266
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAsnGlu 60
Db 1684267 GGCACAAACCCCAATATATGACGCGGTAAAGAGCCGAAACCGATTTATTGTATAAGGAA 1684326
Qy 61 ThrThrLysAenTyLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 1684327 AAACCAAGAAATATCTGACTATCGATCGATTCGCGATTATTAACGAAACAAACAAAGCA 1684386
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 1684387 CTCCTTTCGTAAGATTCTGAAGTCATCAATCTAATCGAGCAGACAGTACAAAGT 1684446
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 1684447 TTAGCGGAAACAGGTGATACGATTCGCGCAAAATTTATAACGCCAACTTAAAGCA 1684506
Qy 121 LysArgValTyrValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
Db 1684507 CAAATGTTGGTWCAGCACCAACTTGGCCAAACCAATGCGATTTTCAATGCCTGTC 1684566
Qy 141 GlyLeuGluValArgGluTyAlaTyTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 1684567 GGTATGACCATTCGTGAATATCGTTATATGATCTGAACGCAAGCCCTTGATGGGAA 1684626
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 1684627 CACTTATTAGAAGATTTAAGCCAAAGCAAGCGCATGTGTGCTTTTACACGGTTGT 1684686
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCGACTGGTATTGACCCCTACTCCAGAAACAATGGCAAGAAATTAGCCGCACTT 1684746
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220
Db 1684747 TCACCTAAAATAGTTGGTGGCCACTCTTGGCTTATCAAGGTTTAGCCCAAGCA 1684806
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTTACGTGCTTTTCAGCAAAACCAACCAAGAATATTAGTG 1684866
Qy 241 AlaSerSerTySerLysAenPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTTCATCTCCGAAAACTTTGGTTTATATATGAACGTTGTGTGCTATTACCCCTT 1684926
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGCAGAAAATGAGAAATGATCACTCATTAACCCCAAGTAAATCAATATTTCGC 1684986
Qy 281 AlaAsnTySerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300

RESULT 9

US-10-981-687-1

; Sequence 1, Application US/10981687
; Publication No. US20050131222A1

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof

; FILE REFERENCE: PB186P2C1D12

; CURRENT APPLICATION NUMBER: US/10/981,687

; CURRENT FILING DATE: 2004-11-05

; PRIOR APPLICATION NUMBER: US 10/158,865

; PRIOR FILING DATE: 2002-06-03

; PRIOR APPLICATION NUMBER: US 09/557,884

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: US 08/476,102

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/426,787

; PRIOR FILING DATE: 1995-04-21

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 1

; LENGTH: 1830121

; TYPE: DNA

; ORGANISM: Haemophilus influenzae

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (4747)..(4747)

; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (9921)..(9921)

; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (10150)..(10150)

; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (29298)..(29298)

; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (36543)..(36543)

; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (36551)..(36551)

; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:

; NAME/KEY: misc_feature

LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)

OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

Alignment Scores:
Pred. No.: 1,27e-138 Length: 1830121
Score: 1325.00 Matches: 248
Percent Similarity: 77.8% Conservative: 60
Best Local Similarity: 62.6% Mismatches: 88
Query Match: 64.8% Indels: 0
DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-981-687-1 (1-1830121)

Qy 1 MetPheGluAenIleThrAlaAapProAlaAapProIleLeuGlyLeuAlaAapLeuPhe 20
Db 1684147 ATGTTTGAACATATCAAGACGGCCAGCCGATCCAATCTTAGGCTTAGCGAAGCAATTT 1684206

Qy 21 ArgAlaAapGluAArgProGlyLysIleAenLeuGlyIleGlyValTyLysAapGluThr 40
Db 1684207 AAATCCGAAACTCGGAATAAATCAATTTGGGTATGGCGTTTATAAGATGCCGAA 1684266

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAenGlu 60
Db 1684267 GGCACAAACCCCAATATGACGCGGTAAAGAGCCGAAAGCAAAACGATTTATTGATAAGGAA 1684326

Qy 61 ThrThrLysAenTyLeuGlyLysIleAapGlyLysIleProGluPheGlyAArgCysThrGlnGlu 80
Db 1684327 AAAACCAAGAAATATCTGACTATCGATGGTATTCGGGATTATAACGAAACAAACAGCA 1684386

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAapLysArgAlaAArgThrAlaGlnThr 100
Db 1684387 CTCCTTTTCGTAAGATCTGAAGTCATCAATCTAATCGAGCAAGACAGTACAAAGT 1684446

Qy 101 ProGlyGlyThrGlyAlaLeuAArgValAlaAlaAapPheLeuAlaLysAenThrSerVal 120
Db 1684447 TTAGCGGAAACAGGTGATACGCAATTCGGCGAGAAATTTATAACGCCAAACTAAAGCA 1684506

Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 1684507 CAAAATGTTGGTWCAGCACCAACCAATTCGGCAACACCAATTCGGATTTTCAATGCTGTC 1684566

Qy 141 GlyLeuGluValAArgGluTyAlaTyTrpAapAlaGluAenHisThrLeuAapPheAap 160
Db 1684567 GGTATGACCATTCGTGAATATCGTTATATGATCTGACGCAAGCCCTTGATGGGAA 1684626

Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAapValValLeuPheHisGlyCys 180
Db 1684627 CACTTATTAGAAGATTTAAGCAAGCAAGCAAGCGATGTGGTCTTTACACGGTTGT 1684686

Qy 181 CysHisAenProThrGlyLysAapProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 1684687 TGCCATAATCCGACTGTTGATGACCCCTACTCCAGAAACATGGCAAGAAATTAGCCGCACTT 1684746

Qy 201 SerValGluLysGlyTrpLeuProLeuPheAapPheAlaTyTrpGlnGlyPheAlaAArgGly 220
Db 1684747 TCACCTAAAATGGTTGGTTGCCACTCTTTGACCTTTGCTTATCAAGTTTAGCCCAACGGA 1684806

Qy 221 LeuGluAapAlaGluGlyLeuAArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 1684807 TTAGATGAAGATGCTTATGGTTTACGTGCTTTTCAGCAACCAACCAAGAAATATTAGTG 1684866

Qy 241 AlaSerSerTySerLysAenPheGlyLeuTyTrpAenGluAArgValGlyAlaCysThrLeu 260
Db 1684867 GCGAGTTCATCTCGAAAAACCTTTGGTTTATATATGAACGTGTGTGGTACATTTACCCCTT 1684926

Qy 261 ValAlaAapSerGluThrValAapArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1684927 GTGCGAGAAATGAGAAATGTCATCAACCTATTAAACCCCAAGTAAATCAATTTATTCGC 1684986

Qy 281 AlaAenTySerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 1684987 GTGCGAGAAATGAGAAATGTCATCAACCTATTAAACCCCAAGTAAATCAATTTATTCGC 1684986

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

Alignment Scores:
Pred. No.: 1,4e-142 Length: 966
Score: 1318.00 Matches: 246
Percent Similarity: 86.3% Conservative: 32
Best Local Similarity: 76.4% Mismatches: 44
Query Match: 64.4% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-23918 (1-966)

Qy 75 GlyArgCysThrGlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAenAapLysArg 94
Db 1 GCGCGGTGATTTCAGNACTGCTTTTGGAAAGTACCAAGTCCATCGCTTACTGAAAACGCT 60

Qy 95 AlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuAArgValAlaAapPheLeu 114
Db 61 GCGCGTACCGTACAAAGCCAGGCGGAAACAGGCGCACTGCGCTACTGCGCTGATTTTATT 120

Qy 115 AlaIleAenThrSerValLysAArgValTrpValSerAenProSerTrpProAenHisLys 134
Db 121 GCGAAGCAAACTAATGCGCAACCGTTTGGATCAGTAACCAACCTGCGCAACCACTATAA 180

Qy 135 SerValPheAenSerAlaGlyLeuGluValAArgGluTyTrpAlaTyTrpAapAlaGluAen 154
Db 181 GTGTGTTTTTCCAGCGCAGGTTTAGAGATCCGCGAATATAACTATTACGATGCAGAAAAA 240

; RESULT 10
; US-10-369-493-23918
; Sequence 23918, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23918
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
; US-10-369-493-23918

Alignment Scores:
Pred. No.: 1,4e-142 Length: 966
Score: 1318.00 Matches: 246
Percent Similarity: 86.3% Conservative: 32
Best Local Similarity: 76.4% Mismatches: 44
Query Match: 64.4% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-23918 (1-966)

Qy 75 GlyArgCysThrGlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAenAapLysArg 94
Db 1 GCGCGGTGATTTCAGNACTGCTTTTGGAAAGTACCAAGTCCATCGCTTACTGAAAACGCT 60

Qy 95 AlaArgThrAlaGlnThrProGlyGlyThrGlyAlaLeuAArgValAlaAapPheLeu 114
Db 61 GCGCGTACCGTACAAAGCCAGGCGGAAACAGGCGCACTGCGCTACTGCGCTGATTTTATT 120

Qy 115 AlaIleAenThrSerValLysAArgValTrpValSerAenProSerTrpProAenHisLys 134
Db 121 GCGAAGCAAACTAATGCGCAACCGTTTGGATCAGTAACCAACCTGCGCAACCACTATAA 180

Qy 135 SerValPheAenSerAlaGlyLeuGluValAArgGluTyTrpAlaTyTrpAapAlaGluAen 154
Db 181 GTGTGTTTTTCCAGCGCAGGTTTAGAGATCCGCGAATATAACTATTACGATGCAGAAAAA 240
```


Qy 302 AlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnAArgileGlnArgMet 321
 Db 907 GATTGGAAGCACAAATGATGCGCAACTCGATGAAATGCGCGCGGTATCAAAGCCATG 966
 Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheile 341
 Db 967 CGCCAAAATTTGTCGAGTTGCTCAAGCCAAAGGTGCAACTCAAGACTTTTGATTTTCATT 1026
 Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
 Db 1027 ATCGACAAACAGCGATGTTCTCTTACGCGGCTTGACTCCCGAACAGTCGACCGCTTAA 1086
 Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
 Db 1087 AAAAAAGAGTTTGCATTTATGCGCTCGCTCCGCGCGCATCAACGTCGCGGCGATTACC 1146
 Qy 382 ProAspAsnMetAlaProLeuLysGluAlaIleValAlaValLeu 396
 Db 1147 GACGACACATCGATTATCTGCGAAAGTATCGTNAAGTACTG 1191

RESULT 12

US-10-915-740A-40
 ; Sequence 40, Application US/10915740A
 ; Publication No. US20050191316A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frazer, Claire M.
 ; APPLICANT: Hickey, Erin
 ; APPLICANT: Peterson, Jeremy
 ; APPLICANT: Tettelin, Hervé
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: Masignani, Vega
 ; APPLICANT: Galeotti, Cesira
 ; APPLICANT: Mora, Manroza
 ; APPLICANT: Ratti, Giulio
 ; APPLICANT: Scarselli, Maria
 ; APPLICANT: Scarlato, Vincenzo
 ; APPLICANT: Rappuoli, Rino
 ; APPLICANT: Pizza, Mariagrazia
 ; APPLICANT: Grandi, Guido
 ; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
 ; FILE REFERENCE: 002441.00090
 ; CURRENT APPLICATION NUMBER: US/10/915,740A
 ; CURRENT FILING DATE: 2004-08-11
 ; PRIOR APPLICATION NUMBER: 09/806,866
 ; PRIOR FILING DATE: 1999-10-08
 ; PRIOR APPLICATION NUMBER: USN 60/103,794
 ; PRIOR FILING DATE: 1998-10-09
 ; PRIOR APPLICATION NUMBER: USN 60/132,068
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/25373
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 1068
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 40
 ; LENGTH: 17381
 ; TYPE: DNA
 ; ORGANISM: Neisseria meningitidis
 US-10-915-740A-40

Alignment Scores:
 Pred. No.: 8,01e-140 Length: 17381
 Score: 1310.00 Matches: 244
 Percent Similarity: 78.0% Conservative: 64
 Best Local Similarity: 61.8% Mismatches: 87
 Query Match: 64.1% Indels: 0
 DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-915-740A-40 (1-17381)

Qy 2 PheGluAsnIleThrAlaIleProAlaAspProfileLeuGlyLeuAlaAspLeuPheArg 21
 Db 8597 TTCAAGCACATCGAAGCGCGCCCGCGGATTCCTCGGTTTGGCGAAGCGTTCAAA 8656

Qy 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThrGly 41
 Db 8657 GCCGAAACCCGCCCGCAAAAGTCAACCTCGCATCGCGCTTTATATAAGACGCGTCCGC 8716
 Qy 42 LysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
 Db 8717 GCACACCCCTCGTCAAGCGCGTCAAGAAAGCGCAAAACCGCTGTGTGGAAGCGAAACC 8776
 Qy 62 ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
 Db 8777 ACCAAAACTACCTACCATCGAGCGGTTCGCGACTACAAACGCGCAACCCAAATCCTG 8836
 Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
 Db 8837 CTGTTCGCAAAAGACCAACGAAATCATCGCCAGCGCTCGCGCCCAAAACAGCGCAAGCCTC 8896
 Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
 Db 8897 GCGCGTACGCGCGCATTTGCGTATTGCGGCGCGAGTTTGCCAAACGCCAGTGTGAACGCGCAA 8956
 Qy 122 ArgValTrrpValSerAsnProSerTrrpProAsnHisLysSerValPheAsnSerAlaGly 141
 Db 8957 ACCATCTGGATTCCAAATCGACTTGGCCCAACACACAGCCCATCGCCAAAGCGGTCTG 9016
 Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
 Db 9017 ATCCAAAGACAAACCTTATCGTTACTATGATGCGCCCAACACACGCTTTGGATTGGAGCGC 9076
 Qy 162 LeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
 Db 9077 ATGATTGAGGACTTGAGCCAAAGCGCAAAAGGCGACATCGCTCTGTCGCGGTCTGTCG 9136
 Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTrrpGlnThrLeuAlaGlnLeuSer 201
 Db 9137 CACAATCTTACCGGCATCGACCTCGCCCAACAAATGGGAACATTTGGCAAACTTTCT 9196
 Qy 202 ValGluLysGlyTrrpLeuProLeuPheAspPheAlaTrrpGlnGlyPheAlaArgGlyLeu 221
 Db 9197 GCCGAAAAAGCGTGGTTGCGCTGTTGACTTTGCTTACCAAGCGCTTCGCGCAATGGTTTG 9256
 Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMethHisLysGluLeuIleValAla 241
 Db 9257 GAAGAAGATGCGCTACCGCTTGGCGGTTCCTTGAACACAAATACAAATGCTGATTGCGC 9316
 Qy 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
 Db 9317 AGCTCTTATTCAAAACTTCGGTATGTACACAGCGCGCTCGCGCGGTTCACCTTTGGTG 9376
 Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
 Db 9377 GCCGAAGATGAAGAAACAGCAGCGCCGCGCCACAGCCAGTCAAAACCATCATCCGTACC 9436
 Qy 282 AsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
 Db 9437 TTGTATTCCAACCCCGCTTCACACGCGTGCAACCAACCATTCGCGTGTGTGTGAAATATGAT 9496
 Qy 302 AlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
 Db 9497 GATTTGAAGCACAAATGGATTGCGCACTCGAATGCGCGCGCGCATCAAGCCCATG 9556
 Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGlyGlyAlaAsnArgAspPheSerPheile 341
 Db 9557 CGCCAAAAATTTGTCGCGTTGCTCAAGCGCAAGGTGCAAGCCAAACATTTTGATTTTCATT 9616
 Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
 Db 9617 ATCAAAACAAACCGGTATGTTCTTTACGCGGTTCGACTCCCGCAACAGTCGACCGCGCTG 9676
 Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
 Db 9677 AAAAACGAGTTTGGCATTTATGCGTCCGCTCCGCGCGCATCAACGTCGCGCGCATTACC 9736

Qy 382 ProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 9737 GACAAACACATCGATTATCTGTGCGAAGCATCGTGAAAGTACTG 9781

RESULT 13

US-10-915-740A-1068/C

; Sequence 1068, Application US/10915740A

; Publication No. US20050191316A1

; GENERAL INFORMATION:

; APPLICANT: Frazer, Claire M.

; APPLICANT: Hickey, Erin

; APPLICANT: Peterson, Jeremy

; APPLICANT: Tettelin, Hervé

; APPLICANT: Venter, J. Craig

; APPLICANT: Massignani, Vega

; APPLICANT: Galeotti, Cesira

; APPLICANT: Mora, Manroza

; APPLICANT: Ratti, Giulio

; APPLICANT: Scarselli, Maria

; APPLICANT: Scarlato, Vincenzo

; APPLICANT: Rappuoli, Rino

; APPLICANT: Pizza, Mariagrazia

; APPLICANT: Grandi, Guido

; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use

; FILE REFERENCE: 002441.00090

; CURRENT APPLICATION NUMBER: US/10/915,740A

; CURRENT FILING DATE: 2004-08-11

; PRIOR APPLICATION NUMBER: 09/806,866

; PRIOR FILING DATE: 1999-10-08

; PRIOR APPLICATION NUMBER: USN 60/103,794

; PRIOR FILING DATE: 1998-10-09

; PRIOR APPLICATION NUMBER: USN 60/132,068

; PRIOR FILING DATE: 1999-04-30

; PRIOR APPLICATION NUMBER: PCT/US99/25373

; PRIOR FILING DATE: 1999-10-08

; NUMBER OF SEQ ID NOS: 1068

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1068

; LENGTH: 2242716

; TYPE: DNA

; ORGANISM: Neisseria meningitidis

US-10-915-740A-1068

Alignment Scores:

Pred. No.: 9,44e-137 Length: 2242716

Score: 1310.00 Matches: 244

Percent Similarity: 78.0% Conservative: 64

Best Local Similarity: 61.8% Mismatches: 87

Query Match: 64.1% Indels: 0

DB: 9 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-915-740A-1068 (1-2242716)

Qy 2 PheGluAenIleThrAlaAlaProAlaAspProfileLeuGlyLeuAlaAspLeuPheArg 21
Db 562972 TTCAGACATCGAAGCGCCCGCCCGCATCGATTCTCGGTTTGGCGAAGCGTTCAA 562913

Qy 22 AlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThrGly 41
Db 562912 GCCGAAACCCCGCCGAAAGATCAACCTCGGCATCGCGGTTTATAAAGACGCGATCCGCGC 562853

Qy 42 LysThrProValThrSerValLysAlaGluGlnTyLysLeuGluAenGluThr 61
Db 562852 CGGACACCCCTCGTCAAGACCGTCAAGAGCCGAAACCGCTGTGGAAGCGAAACC 562793

Qy 62 ThrLysAsnTyLysLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGluLeu 81
Db 562792 ACCAAAACTACCTCACCATCGACGCGGTGCGGACTACACGCGCAACCCCAATCTCG 562733

Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 562732 CTGTTCCGCAAGACACCAATCATCCGACCGCTCGCGCAAAACAGCGCAAGCCTC 562673

Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerValLys 121
Db 562672 GCGCGTACGGCGCATTCGTTATTCGCGCCGAGTTTGCACAAACGCGAGTTGCAACGCGCAA 562613

Qy 122 ArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 562612 ACCATCTGATTTCATCCAGCTTGGCCCAACACAAACGCGCATCGCAACGCGTGGT 562553

Qy 142 LeuGluValArgGluTyAlaTyTrpAspAlaGluAenHisThrLeuAspPheAspAla 161
Db 562552 ATCCAGACAAACCTTATCTGTTACTATGATCGCCCAACACGCGTTTGGATGGGACGCG 562493

Qy 162 LeuLeAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 562492 ATGATTGAGGACTTGGAGCAAGCGCAAAAGCGCACATCGTCTCTGCTCAGCGGCTGCTGC 562433

Qy 182 HisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSer 201
Db 562432 CACAATCTTACCGGCATCGACCTTACGCGCAACAAATGGAAACCTTTGGCAAACTTTCT 562373

Qy 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArgGlyLeu 221
Db 562372 GCCGAAAGAGCTGGTTGCCGCTGTTGACTTTGGCTACCAAGGCTTCGCAATGGTTG 562313

Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuValAla 241
Db 562312 GAAGAAGATGCTTACGGCTTGGCGGTGTTCTTGAACAACAATACAGAATTCGTGATTGCC 562253

Qy 242 SerSerTyLysAsnPheGlyLeuTyTrpAsnGluArgValGlyAlaCysThrLeuVal 261
Db 562252 AGCTCTTATTCCAAAACCTTCGGTATGTACAAACGCGCGCTCGCGCGCTTCACATTTGGTG 562193

Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db 562192 CCCGAAAGATGAAGAACAGAGCGCGCCGACAGCGCAAGTCAAAACCATCATCGGTACC 562133

Qy 282 AsnTyLysAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAep 301
Db 562132 TTGTATTCCAACCGCGCTTCACCGTTCGCAACACCATTCGCTGGTGTGAAAAATGAT 562073

Qy 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 562072 GATTTGAAGCACCAATGGAATTCGCGAACTCGATGAAATGCGCGCGCATCAAAAGCCATG 562013

Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheIle 341
Db 562012 CGCCAAAAATTTGCGGGTTCCTCAAGCCCAAGGTGCAAGCCAAACTTTGATTTCAAT 561953

Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 561952 ATCAACAAACCGGTATGTTCTCTTTCAGCGGCTTGACTCCCGAACAAGTCGACCGCCTG 561893

Qy 362 ArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 561892 AAAACGAGTTTGCATTTATCGCTCCGCTCCGCGCGCATCAACGTCGCGCGCATTACC 561833

Qy 382 ProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 561832 GACAAACATCGATTATCTGTGCGAAGCATCGTGAAAGTACTG 561788

RESULT 14

US-10-369-493-35511

; Sequence 35511, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35511
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35511

Alignment Scores:
Pred. No.: 3.89e-106 Length: 1185
Score: 1005.50 Matches: 199
Percent Similarity: 65.9% Conservativeness: 62
Best Local Similarity: 50.3% Mismatches: 132
Query Match: 49.2% Indels: 3
DB: 6 Gaps: 1

US-10-673-786A-2 (1-396) x US-10-369-493-35511 (1-1185)

Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 1 ATGTTGAGACCTGACAGCCAGCTCCGCCGACAAAGATCTGGCCCTCATCGGCTCTAT 60
Qy 21 ArgAlaAaspGluAArgProGlyLysIleAenLeuGlyLeuGlyValTyrLysAaspGluThr 40
Db 61 CGCGCCGATCCGCTCCCAACAGAGTGCACCTCGGCGTCTACAAAGGATCGCGAC 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAGACCCAGATGATCGCGCGCTGCGTGGAGCTGAAAGAGCGGCTGCTGAACAGCGAC 180
Qy 61 ThrThrLysAenTyrLeuGlyLysIleAaspGlyProGluPheGlyAArgCysThrGlnGlu 80
Db 181 GACACCAAGACCTATCTCGGCTCGCGCGACACCGGCTTCAACGGCGTATGCGCCAG 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuAenAaspLysArgAlaAArgThrAlaGlnThr 100
Db 241 CTCGCTTCGGCCCGCGCGGACATG-----ACGGCATTTCGGCGGCCCGACGGC 291
Qy 101 ProGlyThrGlyAlaLeuAArgValAlaAaspPheLeuAlaLysAenThrSerVal 120
Db 292 CCTGGCGCTCGGCTGCGCTTCTGCTGGCGAGCTGCTCAAGCGGACCCGCTTCGGAC 351
Qy 121 LysArgValTyrValSerAenProSerTyrProAenHisLysSerValPheAenSerAla 140
Db 352 GCCACGATCTGGCTGCGACCACTTTCGCGGACCATATGCGCGTGTCTGCTGCGCTGCC 411
Qy 141 GlyLeuGluValAArgGlyTyrAlaTyrTyrAaspAlaGluAenHisThrLeuAaspPheAasp 160
Db 412 GGCCTGCAGATCCGCAATATCCCTATTTCGACGCGGCTTTCGCGTGTCTGCTGCTGCAC 471
Qy 161 AlaLeuLeuAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 472 GACATGTGCGCGGCTTTCGCGACGCCAAGAGCGGCGAGCTGCTGCTGCTGCTGCTGCTG 531
Qy 181 CysHisAenProThrGlyLysIleAenProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 532 TGCCACAAATCCACCGCGGCAATCTGATGCGCGGAGTGGGAGCGGCTGCGGCTGCTGCTG 591
Qy 201 SerValGluLysGlyTyrPheLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAArgGly 220
Db 592 ATGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
Qy 221 LeuGluAaspAlaGluGlyLeuAArgAlaPheAlaAmetHisLysGluLeuLeuVal 240
Db 652 CTCGAGCGCGAGCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluAArgValGlyAlaCysThrLeu 260
Db 712 GCCTCGAGCTGCTGCAAGAAATTTCCCGCTTACCGCGACCGCTGCTGCGCGGACGATGGTT 771

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31184
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-10-369-493-31184

Alignment Scores:
Pred. No.: 4.95e-105 Length: 1188
Score: 996.00 Matches: 191
Percent Similarity: 65.2% Conservativeness: 67
Best Local Similarity: 48.2% Mismatches: 138
Query Match: 48.7% Indels: 0
DB: 6 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-31184 (1-1188)

Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 1 CTGTTCTCCGCGCTCGAATCTGCTCCCGGACCCGATTCCTGGGCTTGAACGAAGCCTTC 60
Qy 21 ArgAlaAaspGluAArgProGlyLysIleAenLeuGlyLysValTyrLysAaspGluThr 40
Db 61 AATGCGCATACGCGCGACCCACCAAGGTCACCTTGGCGTGGCGTGTACTTCAATGAAGAA 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAGATTCCTGCTGCGCGCTGCGCGACGCAAGAAAGCCCGCGCTGATGCGCGCG 180
```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 16:06:55 ; Search time 442 Seconds
(without alignments)
2089.040 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAPADPIGLDLF.....VAGTMDNAPLCEAIVAVL 396

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abs/ABSSWEB_spool/US10673786/runat_14032006_135624_22906/app_query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=rnpbnp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss03h
-USER=US10673786 -CGCN_1_1_431 @runat_14032006_135624_22906 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:

1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2011	98.3	1191	12	US-11-114-922-75
2	1317	64.4	1269	8	US-10-467-657-6535
3	842	41.2	1191	8	US-10-467-657-1059
4	831.5	40.7	1194	8	US-10-979-821-31
					Sequence 75, Appl
					Sequence 6535, Ap
					Sequence 1059, Ap
					Sequence 31, Appl

5	831.5	40.7	1194	12	US-11-114-922-31	Sequence 31, Appl	
6	818	40.0	1260	8	US-10-979-821-5	Sequence 5, Appl	
7	818	40.0	1260	12	US-11-114-922-5	Sequence 5, Appl	
8	815	39.9	1170	8	US-10-979-821-1	Sequence 1, Appl	
9	815	39.9	1170	12	US-11-114-922-1	Sequence 1, Appl	
10	810	39.6	1260	8	US-10-979-821-3	Sequence 3, Appl	
11	810	39.6	1260	12	US-11-114-922-3	Sequence 3, Appl	
12	781	38.2	2343	12	US-11-136-527-2278	Sequence 2278, Ap	
13	727.5	35.6	1251	7	US-10-932-182A-82350	Sequence 82350, A	
14	727.5	35.6	1251	7	US-10-932-182A-82350	Sequence 82350, A	
15	721.5	35.3	1389	7	US-10-932-182A-4671	Sequence 4671, Ap	
16	721.5	35.3	1389	7	US-10-932-182A-4671	Sequence 4671, Ap	
17	632.5	30.9	1239	8	US-10-979-821-7	Sequence 7, Appl	
18	632.5	30.9	1239	12	US-11-114-922-7	Sequence 7, Appl	
19	616.5	30.1	1356	7	US-10-932-182A-1541	Sequence 1541, Ap	
20	616.5	30.1	1356	7	US-10-932-182A-1541	Sequence 1541, Ap	
21	597.5	29.2	1356	7	US-10-932-182A-77511	Sequence 77511, A	
22	597.5	29.2	1356	7	US-10-932-182A-77511	Sequence 77511, A	
23	348	17.0	560	6	US-09-925-065A-479173	Sequence 479173, A	
24	348	17.0	560	6	US-09-925-065A-479173	Sequence 479173, A	
25	281	13.7	1400	12	US-11-136-527-6374	Sequence 6374, Ap	
26	191	9.3	594	6	US-09-925-065A-739968	Sequence 739968, A	
27	153.5	7.5	1290	8	US-10-793-626-3173	Sequence 3173, Ap	
28	153.5	7.5	2883	8	US-10-793-626-3173	Sequence 3173, Ap	
C	29	152.5	7.5	2631	8	US-10-793-626-3760	Sequence 3760, Ap
	30	124.5	6.1	1230	12	US-11-055-822-115	Sequence 115, App
	31	124.5	6.1	1230	12	US-11-055-822-485	Sequence 485, App
32	119	5.8	1119	7	US-10-932-182A-3622	Sequence 3622, Ap	
33	119	5.8	1119	7	US-10-932-182A-3622	Sequence 3622, Ap	
C	34	112.5	5.5	922	8	US-10-750-185-38121	Sequence 38121, A
	35	112.5	5.5	922	8	US-10-750-623-38121	Sequence 38121, A
	36	111	5.4	1356	8	US-10-467-657-1849	Sequence 1849, Ap
37	109.5	5.4	1161	7	US-10-932-182A-3002	Sequence 3002, Ap	
38	109.5	5.4	1161	7	US-10-932-182A-3002	Sequence 3002, Ap	
39	108	5.3	2529	12	US-11-136-527-3251	Sequence 3251, Ap	
40	105.5	5.2	1194	9	US-11-232-405A-41	Sequence 41, Appl	
C	41	103.5	5.1	4235	8	US-10-793-626-4041	Sequence 4041, Ap
	42	103	5.0	5190	12	US-11-091-883-31	Sequence 31, Appl
	43	103	5.0	5190	12	US-11-091-883-177	Sequence 177, Appl
44	99	4.8	1182	8	US-10-979-821-9	Sequence 9, Appl	
45	99	4.8	1182	12	US-11-114-922-9	Sequence 9, Appl	

ALIGNMENTS

RESULT 1
US-11-114-922-75
; Sequence 75, Application US/1114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 75
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-11-114-922-75

Alignment Scores:
Pred. No.: 2.76e-209
Score: 2011.00
Percent Similarity: 99.2%
Length: 1191
Matches: 390
Conservative: 3


```

Db 448 ACCATCTGGATTCCAAATCCGACCTGGCCCAACCAACAGCCATTCGCCAAACGGCTCGGT 507
Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAspAla 161
Db 508 ATCCAGACAAACCTTATCGCTACTATGATGCGCCCAACACACTGCTTGGATTGGAGCGGC 567
Qy 162 LeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysCys 181
Db 568 ATGATTGAAGATTGAACCAAGCCAAAGGCGACATCGTCTGCTGACGGCTGCTGC 627
Qy 182 HisAsnProThrGlyLeuAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeuSer 201
Db 628 CACAACCCCTACCGGTATGACCTACGCCCGCAACCAATGGAACTTTAGCAAACTTCT 687
Qy 202 ValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
Db 688 GCCGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
Qy 222 GluGluAspAlaGluGluGluAlaGluAlaPheAlaAlaMetHisLysGluLeuLeuValAla 241
Db 748 GAAGAAGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Qy 242 SerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeuVal 261
Db 808 AGCTCTTATCCAAAACCTTCGCGATGTACCAACGAGCGTGTGCGGCACTTCACTTTGGTG 867
Qy 262 AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaAlaAlaAla 281
Db 868 GCCGAGATGAAGAAACAGCAGCCCGCCGCGCCACACCAATCAAAACCATCATCCGTACC 927
Qy 282 AsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsnAsp 301
Db 928 TTGTATTCCAAACCGCTTCACATGTCGACACACCATTCGCTGCTGTGTAATAATGAT 987
Qy 302 AlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArgMet 321
Db 988 GATTGGAAGCACAATGATGCTGCGAACTTGACGAATGCGGCGCGCATCAAGCCCATG 1047
Qy 322 ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPheIle 341
Db 1048 CCCCAAAATTTGTCAGTTGCTCAAGGCCAAAGTTCAACCAAGACTTGAATTTAT 1107
Qy 342 IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 1108 ATTGAACAAACGGTATGCTCTCTTTCAGCGCTTGACTCCCGAACAAGTTGACCGTTG 1167
Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
Db 1168 AAAACGAGTTTGCCTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIleValAlaVal 395
Db 1228 GACGACAAATCGATTATCTGTGTGAAGCATCGTGAAGAATA 1269

```

RESULT 3

```

; Sequence 1059, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1059

```

```

; LENGTH: 1191.
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1059

Alignment Scores:
Pred. No.: 1.03e-81 Length: 1191
Score: 842.00 Matches: 165
Percent Similarity: 58.8% Conservative: 68
Best Local Similarity: 41.7% Mismatches: 163
Query Match: 41.2% Indels: 0
DB: 8 Gaps: 0

US-10-673-786a-2 (1-396) x US-10-467-657-1059 (1-1191)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTACCGACACATCGAATATCTATCCCGCGACCCGATTTTGAGTTTGGTTCGAAACCTTC 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 AAAAACGACCCGCGCCGCGGAAAAAGTCAATTTGACATAGGCATTTATTTTCGACGACGAA 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 121 GCGAATGCGCGTATTTGGAATCCGTAAGCGCGTGCAGAACCCGCGCGCGCGCCCA 180
Qy 61 ThrThrLysAsnTyrLeuGlyLysAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 GCGCGCTCGCGCTACTCCGATGAGGGTTTGGACGTTTACCGAGCGCGGTGCGAGCAT 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 TTGTTGTTGCGCAAGGCAACCCCGCTCGCGCAGGAGCGCATCGTTACCGTACAGACT 300
Qy 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CTGGCGCGCTCGGCGCGCTCAAGTGGGGCGGATTTCTGCACCGCTGTTTCCCGAA 360
Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 361 GCGCGCGCTACGTCAGCGATCCGATTTGGGCAACCATCGCGCGCATTTTGAAGCGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGTTCGAGGTGCGCACTTACCATATTTACGACCTGCGCCTGCTGCGCGTGAATTCGAC 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GAAATGACCGGCTTTTCAACACCTGCGCCGCAACAGCGTCTCATCTCCATCCCTGCG 540
Qy 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 541 TGCCACAAACCCGCGCGCGTATGTCGAAACCGCAATGGGACGAAGTGTTCGAGATC 600
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 ATCAAAACGCGCAAACTGATTCGTTTATGACATCGCTACCAAGGCTTCGCGCGCGAT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 TTGGACGCGATGCTACGCGCTCGCAACGCGGTGGAATGGAATTTGCGCGTGTGTTGTC 720
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 721 AGCAATTCGTTTCAAAAAACCTGCTGCTACGCGCAACGCGTCCGCGGCTTAAGCGGTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTTCGCCCAATAAAGAGAGCGGATTTGGTTGTCGACAACTCAATTTCCCGTCCGCG 840
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300

```

Db 841 CGATCTACTCCAGCCGCCGCGCGATATATCGCCGCCGAGTGTGAACAGC 900
Qy 301 AspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 TCCGAACCTTACGCTTGTGGCAACAGAACTTATATGATGCGCGACCATCCGTGCG 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 961 ATGGGCGAGAACTTACGGCGTATTGACCGCCGAAATCCCGACCGCGATTTCACTTAT 1020
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 TTCAATCAACAGCGCGAAATGTTCCGCTACACGGGGTTGACGGTGGGCAAGTCCGCGAG 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 TTGGCGCAGCAATTTGCGCTTACTCTGCTGCGGAGGATGCGTCCGCGGCTG 1140
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 AATACATCAAAATATCACTTATGTCGCGATGCGCTTGCAGAAAGTGTG 1188

RESULT 4

US-10-979-821-31

; Sequence 31, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA W.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: PRODUCTION OF MONATIN AND ITS PRECURSORS
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 31
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Escherichia coli

US-10-979-821-31

Alignment Scores:
Pred. No.: 1.45e-80 Length: 1194
Score: 831.50 Matches: 167
Percent Similarity: 60.7% Conservative: 74
Best Local Similarity: 42.1% Mismatches: 155
Query Match: 40.7% Indels: 1
DB: 8 Gaps: 1

US-10-673-786A-2 (1-396) x US-10-979-821-31 (1-1194)

Qy 1 MetPheGluAsnIleThrAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 GTGTTTCAAAAGTTGAGCCCTACGCTGGCGACCCGATTTCTACGCTTATGGACGCTTT 60
Qy 21 ArgAlaAspGluArgProGlyIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 AAAGAAGACCCCTCGCAGCGACAAAGTGAATTAAGTATCGGTCTGTACTACAACGAAGAC 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeu---LeuGluAsn 59
Db 121 GGAATTATTCACAACTGCAACCGCTGGCGAGCGGAAGCGCGCTTGAATGCGCAGCCT 180

Qy 60 GluThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db 181 CATGGCGCTTCGCTTTTATTCAGCATGGAAGGGCTTAACATGCTATGCCATGCCATTCGCG 240
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGln 99
Db 241 CCCTGCTGTTTGGTGGGACCATTCGGTACTGAAACAACAGCGCGTAGCAACCATTCAA 300
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
Db 301 ACCCTGGCGCTCCGGGCAATGAAAGTGGCGCGGATTTCTGAAACCGTACTTCCCG 360
Qy 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
Db 361 GAATCAGCGCTCTGGTTCAGCGATCTACCTCGGAAACCGCGTAGCAATATTCGCGCGG 420
Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPhe 159
Db 421 GCTGGATTCGAAGTGAGTACTTACCTCGTATGACGAAGCGACTAACGGCGTGGCTTT 480
Qy 160 AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 481 AATGACCTGTGGCGAGCTGAAACAATACCTGCCCGCAGTATTGTTGCTGCATCCA 540
Qy 180 CysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGln 199
Db 541 TGTTCACCAACCAACCGGTGCGATCTCACTAAATGATCAGTGGGATGCGGTGATTGAA 600
Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
Db 601 ATTCTCAAGCCCGCGAGCTTATTCCATTCCTCGATATTCCTATCAAGGATTTGGTGCC 660
Qy 220 GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIle 239
Db 661 GGTATGGAAGAGATGCTACGCTATTCGCCCATTCGCCAGCGCTGATACCGCTCTG 720
Qy 240 ValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThr 259
Db 721 GTGAGCAATTCGTTCTCGMAAATTTTCTCCCTTTACGGCAGCGCGTGGCGGACTTCT 780
Qy 260 LeuValAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIle 279
Db 781 GTTATGTGAAGATCCGAAGCCGCTGCGCGCTACTGGGCGCAATTTGAAAGCAACAGTT 840
Qy 280 ArgAlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSer 299
Db 841 CGCGCAACTACTCTCAGCCCGCGCAATTTGGTGGCAGGTGGTGGCTGCAGTCTGAAT 900
Qy 300 AsnAspAlaLeuArgAlaIleTrrpGluGlnGluLeuThrAspMetArgGlnArgIleGln 319
Db 901 GACGAGCATTTGAAGCCAGCTGGCTGGCGGAAGTAGAAGATGCGTACTCGCATTTCTG 960
Qy 320 ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSer 339
Db 961 GCAATGCGTCAGGAATTTGGTGAAGAGATTTAAGCACAGAGATGCCAGAACCAATTCGAT 1020
Qy 340 PheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu 359
Db 1021 TATCTGCTTAATCAGCGCGCATGTTCAAGTTATACCGGTTTAAAGTGGCGCTCAGTTGAC 1080
Qy 360 ArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGly 379
Db 1081 CGACTACGTGAAGAAATTTGGTGTCTATCTCATCCAGCGCGTGCATGTGTGTCGCGGG 1140
Qy 380 MetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 TTAATACCGCAAAATGTCAACCGTGTGGCAAGCGCTTGTGCGGTGATG 1191

RESULT 5

US-11-114-922-31

; Sequence 31, Application US/11114922

; Publication No. US20050282260A1

Qy	220	GlyLeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuLeu	239
Db	661	GGTATGGAAGAGGATGCTACCGCTATTTCGGCCATTGCCAGCGCTGGATTACCCGCTCTG	720
Qy	240	ValAlaSerSerTySerLysAsnPheGlyLeuTyRAsnGluArgValGlyAlaCysThr	259
Db	721	GTGAGCAATTCGTTCTCGAAATTTTCTCCCTTTACGGCAGCGCGTCGGCGGACTTTCT	780
Qy	260	LeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIle	279
Db	781	GTTATGTGTGAAGATGCCGAAGCGCTGGCGGGCTACTGGGGCAATTGAAAGCAACAGTT	840
Qy	280	ArgAlaAsnTyRSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSer	299
Db	841	CGCGCAACTACTCCAGCGCGCGAATTTTGGTGGCGAGGTGGTGGCTGCAGTGTCAAT	900
Qy	300	AsnAspAlaLeuArgAlaIleTTrGluGlnGluLeuThrAspMetArgGlnArgIleGln	319
Db	901	GAGCAGCACTTGAAGCCACCTGGCTGGCGGAGTAGAAGAGATGCGTATCTCGCATTTCTG	960
Qy	320	ArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSer	339
Db	961	GCAATGGCTCAGAAATGGTGAAGGTATTAAACACAGAGATGCCAGAACCAATTTCCAT	1020
Qy	340	PheIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeu	359
Db	1021	TATCTGCTTAATCAGCGCGCATGTTTCAGTTATACCGGTTTAAAGTCGCGCTCAGGTTGAC	1080
Qy	360	ArgLeuArgGluGluPheGlyValTyRAlaValAlaSerGlyArgValAsnValAlaGly	379
Db	1081	CGACTACGTGAAGAAATTTGGTGTCTATCTCATCGCCAGCGGTGCGATGTGTGTCGCCGG	1140
Qy	380	MetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
Db	1141	TTAAATACGCCAAATGTACAACTGTGGCAAGCGGTTTCTGCGGTGATG	1191
RESULT 6			
US-10-979-821-5			
; Sequence 5, Application US/10979821			
; Publication No. US20050244937A1			
; GENERAL INFORMATION:			
; APPLICANT: ABRAHAM, TIMOTHY W.			
; APPLICANT: CAMERON, DOUGLAS C.			
; APPLICANT: HICKS, PAULA M.			
; APPLICANT: MCFARLAN, SARA C.			
; APPLICANT: MILLIS, JIM			
; APPLICANT: ROSAZZA, JACK			
; APPLICANT: ZHAO, LISHAN			
; APPLICANT: WEINER, DAVID P.			
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE			
; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS			
; FILE REFERENCE: 023829-0390			
; CURRENT APPLICATION NUMBER: US/10/979, 821			
; CURRENT FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: 10/422,366			
; PRIOR FILING DATE: 2003-04-23			
; PRIOR APPLICATION NUMBER: 60/374, 831			
; PRIOR FILING DATE: 2002-04-23			
; NUMBER OF SEQ ID NOS: 77			
; SOFTWARE: PatentIn Ver. 3.3			
; SEQ ID NO 5			
; LENGTH: 1260			
; TYPE: DNA			
; ORGANISM: Rhodobacter sphaeroides			
US-10-979-821-5			
Alignment Scores:			
Pred. No.: 4 61e-79			
Score: 818.00			
Percent Similarity: 58.1%			
Best Local Similarity: 42.9%			
Query Match: 10			
Length: 1260			
Matches: 171			
Conservative: 61			
Mismatch: 157			
Indels: 10			

```
DB: 8 Gaps: 5
US-10-673-786A-2 (1-396) x US-10-979-821-5 (1-1260)
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 76 ATGTGACCGCCCTGAAGCGCGACCGCGGACCAAGATCTCTGCAACTGATCCAGATGTTT 135
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 136 CGCAGAGATGCGCGCGGACNAGATCGATCTGGCGTGGCGTCTACAGAGACCGGACC 195
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 196 GGGCTACCCCGGTATCGCGCGCGTGAAGCGCGCGGACCGCGGAGCGGCTCTGGAGTGCAG 255
Qy 61 ThrThrLysAsnTyrLeuGlyLysIleAspGlyLysProGluPheGlyArgCysThrGlnGlu 80
Db 256 ACCACCAAGACCTACACCGGCTTGGCGCGGAGCGCCCTACAAATGCGCGATGGCGAAG 315
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 316 CTGATCTCTC-----GCAGGCGCGTCCCGCGCGGCGCGGCTCGGTCCGCCACC 366
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 367 CCGCGCGCACCGCGCGGTGCGTCCAGCGCTCGAGCTGATCCGCATGGGCTCGGCCGACC 426
Qy 121 LysArgValTrpValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
Db 427 GCCACTGTCTGGATCTCGAAGACCGGACCTGGCGCGGACCACTCTGTCGATCGTGAATATCTC 486
Qy 141 GlyLeuGluValArgGlyTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 487 GGCATCCGATGCGGGATACCGCTATTTCAGCCCGGAGACCGCGCGCTCGATGCCGAG 546
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 547 GGCCTTGATGGAGGATCTGGCCCGAGTGAAGCGCGGCGACGCTGTGCTGCTGCACGGTGC 606
Qy 181 CysHisAsnProThrGlyLysAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 607 TGCCACAAACCGCGCGCGCAACCGGTCAGTGGCTGCGGCGGCTCGCGGAGC 666
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 667 CTGGCCCGGACAGCGCGGTGCGCTGATCGACTCGCTCTATCAGGCTTCGCGGACGGG 726
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 727 CTCGAGATGGATGCGCGCGGACGCGCTTCTGGCCACAGACTGCGCGGAGGTGCTGATC 786
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 787 GCGGCTCTCTGCTCGAAGAACTTCGGCATCTACCGGAGCGAAGCGGCGATCTCTGATCGCC 846
Qy 261 Val-----AlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
Db 847 ATCGGCGAGGCGCGCGCGGCGGACCGGTGAG-----GCCAACCTCAACTTCTCTG 897
Qy 279 IleArgAlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 898 AACCGGACAGAACTACTCTTCCCGCGGACCATGGCGCGGCTCGTGCATCATGATCTCTC 957
Qy 299 SerAsnAspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIle 318
Db 958 GAGGACGAGACGCTGAGCGCGGACTGGAAGCGGAGCGGAACTGAGGAGGTGCGGCTCAACATG 1017
Qy 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGln--GluLysGlyAlaAsnArgAsp 337
Db 1018 CTGACGCTGGCGGCCCGAGCTTGGCGGTCGCGTGGCGGCGGAGACCGGCTCGAACCGC--- 1074
Qy 338 PheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGln 357
```

```
Db 1075 TTCGGCTTCGTGGCCGAGCATCGCGCATGTTCTCGCGCTCGGATCACGCCCGCCGAG 1134
Qy 358 ValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnVal 377
Db 1135 GTGAGCGGCTGCGGACCGGACGACGCGGTCTACATGTTGGCGGATTCGCGGCTGAACATC 1194
Qy 378 AlaGlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1195 GCGGCGGTGAACACCGGACGACCGTGGCGGTCTGGCGCGCGGTGGCCAAAGTGCTG 1251
RESULT 7
US-11-114-922-5
; Sequence 5, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; PRIOR FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Rhodobacter sphaeroides
US-11-114-922-5
Alignment Scores:
Pred. No.: 4,61e-79 Length: 1260
Score: 818.00 Matches: 171
Percent Similarity: 58.1% Conservative: 61
Best Local Similarity: 42.9% Mismatches: 157
Query Match: 40.0% Indels: 10
DB: 12 Gaps: 5
US-10-673-786A-2 (1-396) x US-11-114-922-5 (1-1260)
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 76 ATGCTACCGCCCTGAAGCGCGGACCGCGGACCAAGATCTCTGCAACTGATCCAGATGTTT 135
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 136 CGCAGAGATGCGCGCGGACNAGATCGATCTGGCGTGGCGTCTACAGAGACCGGACC 195
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 196 GGGCTACCCCGGTATCGCGCGCGTGAAGCGCGCGGACCGCGGAGCGGCTCTGGAGTGCAG 255
Qy 61 ThrThrLysAsnTyrLeuGlyLysIleAspGlyLysProGluPheGlyArgCysThrGlnGlu 80
Db 256 ACCACCAAGACCTACACCGGCTTGGCGCGGAGCGCCCTACAAATGCGCGATGGCGAAG 315
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 316 CTGATCTCTC-----GCAGGCGCGTCCCGCGCGGCGCGGCTCGGTCCGCCACC 366
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 367 CCGCGCGCACCGCGCGGTGCGTCCAGCGCTCGAGCTGATCCGCATGGGCTCGGCCGACC 426
Qy 121 LysArgValTrpValSerAsnProSerTyrProAsnHisLysSerValPheAsnSerAla 140
Db 427 GCCACTGTCTGGATCTCGAAGACCGGACCTGGCGCGGACCACTCTGTCGATCGTGAATATCTC 486
```



```

Db      769  CGGACCACTCGACTCGCTCGCGGACAGGCTGCGCTCAAACTCGCGGGCTCGCACGC 828
Qy      281  AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn 300
Db      829  ACCAGCTATTCCATCCCGCGGATCACGGCGAGCGCTGTCGGACGATCTTCACGAC 888
Qy      301  AspAlaLeuArgAlaIleTrrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
Db      889  CCGGAACCTCAGCGCGACTGGACGGAGGAGCTCGAGACGATCGCGCTCAGGATCAGGGC 948
Qy      321  MetArgGlnLeuPheValAsnThrLeu-----GlnGluLysGlyValaAsnArg 336
Db      949  CTCGGCGGCTCGCTTCGCGAGGACTCGCCACCGCTCGCAGAGCCTCGCGCGCATCGCC 1008
Qy      337  AspPheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGlu 356
Db      1009  GAT-----CAGGAGGCGATGTTCTCCATGCTGCCGCTTCGAGCG 1050
Qy      357  GlnValLeuArgLeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsn 376
Db      1051  GAGGTTATGCGGCTCAGGACCGGACGACGCGCATCTATATGCGGCGATCGCGCGCATCAAC 1110
Qy      377  ValAlaGlyMet 380
Db      1111  ATCGCGGGCTG 1122

```

RESULT 9

US-11-114-922-1

; Sequence 1, Application US/11114922

; Publication No. US20050282260A1

; GENERAL INFORMATION:

; APPLICANT: MCFARLAN, SARA C.

; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE

; FILE REFERENCE: PRODUCTION OF MONATIN AND ITS PRECURSORS

; CURRENT FILING DATE: 2005-04-26

; PRIOR APPLICATION NUMBER: 10/422,366

; PRIOR FILING DATE: 2003-04-23

; PRIOR APPLICATION NUMBER: 60/374,831

; NUMBER OF SEQ ID NOS: 91

; SOFTWARE: PatentIn Ver. 3.3

; SEQ ID NO 1

; LENGTH: 1170

; TYPE: DNA

; ORGANISM: Sinorhizobium meliloti

US-11-114-922-1

Alignment Scores:

Pred. No.:	8,938-79	Length:	1170
Score:	815.00	Matches:	170
Percent Similarity:	60.7%	Conservative:	63
Best Local Similarity:	44.3%	Mismatches:	137
Query Match:	39.9%	Indels:	14
DB:	12	Gaps:	4

US-10-673-786A-2 (1-396) x US-11-114-922-1 (1-1170)

```

Qy      1  MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db      1  ATGTTGAGCGCTCGCCCGCCACCGACGATCCCTTGCTTTTCTGATCGGCTGTTTC 60
Qy      21  ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db      61  AGGAAGGATGAGCGCGCGGAAAGTCGATCTCGCGCTAGGAGTCTATCGCGACGAGACC 120
Qy      41  GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db      121  GGACGACGCGCATCTTCGGGCGCTCAAGCGCGGAAAGCGGCTTCTCGAACAACAG 180
Qy      61  ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80

```

RESULT 10

US-10-979-821-3

; Sequence 3, Application US/10979821

; Publication No. US20050244937A1

; GENERAL INFORMATION:

; APPLICANT: ABRAHAM, TIMOTHY W.


```

US-10-673-786A-2 (1-396) x US-11-136-527-2 (1-2343)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 76 ATGTCAGCGCCCTAGACCGCGAGCCCGGAGCAAGATCCTGCAACTGATCCAGATGTC 135
Qy 21 ArgAlaAaspGluAargProGlyLysIleAenLeuGlyIleGlyValTyrLysAaspGluThr 40
Db 136 CGCAGGATGCGCGCGCGACAAAGATCGATCTGGCGGTGGCGGTCTACAAGGACCCGAC 195
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 196 GGGCTACCCCGCTATGCGGCGCTGAAGCGCGCGAGAGCGGCTCTGGAGGTGCGAG 255
Qy 61 ThrThrLysAenTyrLeuGlyLysAaspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 256 ACCACCAAGACCTACACCGCCTCCGCGAGCGCGCTCAATAGTCGCGATGCGGAAG 315
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGlnThr 100
Db 316 CTGATCCTC-----GCGGCGCGGTCCGCGCGACCGCGGTGGCTCGGTCGCCAC 366
Qy 101 ProGlyGlyThrGlyAlaLeuAargValAlaAaspPheLeuAlaLysAenThrSerVal 120
Db 367 CCGCGCGACCGCGCGCGGTGGTTCAGCGCTCGAGCTCGAGTCCGCGATGCGCTCGCCGAG 426
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 427 GCCACCGCTGATCTCGAACCGACCTGCGCGAACCATCTGTCGATCGTGAATATCTC 486
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAaspAlaGluAenHisThrLeuAaspPheAasp 160
Db 487 GGCATCCCGATCGCGGAATACCGTATTTCACGCGCGAGACCGCGCGCTCGATGCGCGAG 546
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
Db 547 GGCATGATGGAGGATCTGCCGAGTGAAGCGCGCGCGGTGGTCTGCTCGACGCGTGC 606
Qy 181 CysHisAenProThrGlyLysAaspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 607 TGCCACACCGACCGCGCGCAACCGACCGGTGCGAGTGGTGGCCATCTGCGAGAGC 666
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAargGly 220
Db 667 CTGCGCGCGAGCGCGCGGTGCGCTGATCGACCTCGCTATACAGGCTTCGCGACGCG 726
Qy 221 LeuGluGluAaspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 727 CTCGAGATGGATGCGCGCGCGAGCGGTCTTGGCCACCGAGCTGCCCGAGGTGCTGATC 786
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluAargValGlyAlaCysThrLeu 260
Db 787 GCGGCTCTCTCGAAGAACTTCGGCATCTACCGCGAGCGCACCGGGATCTGATCGCC 846
Qy 261 Val-----AlaAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAla 278
Db 847 ATCGCGAGCGCGCGCGCGCGACGCGTGCAG-----GCCAACCTCAACTTCCTG 897
Qy 279 IleArgAlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 898 AACCGCGAGAACTACTCTCTCCGCGCGACCATGCGCGCGCGCTGTCGATGATCTC 957
Qy 299 SerAenAaspAlaLeuArgAlaIleTrpGluGlnLeuThrAaspMetArgGlnArgIle 318
Db 958 GAGGACGAGACGCTGAGCGCGACCTGGAGCGCGGNACTCGAGGAGTCCGCTCAACATG 1017
Qy 319 GlnArgMetArgGlnLeuPheValAenThrLeuGln---GluLysGlyAlaAenAargAasp 337
Db 1018 CTGACACTGCGCGCGCGCTGCGCGTTCGAGCGCGAGACCGGCTCGAAACGCG--- 1074
Qy 338 PheSerPheIleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGln 357
Db 1075 TTCGGCTTCGTGGCGCGAGCATCGCGCGATGTTCTCGCGCTCGGGATCACCGCGCGAG 1134

```

```

Qy 358 ValLeuArgLeuAargGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenVal 377
Db 1135 GTGAGCGGCTGCGAGACCGAGCACGGGTCTACATGTTGGCGGATTCGCGCTGAACATC 1194
Qy 378 AlaGlyMetThrProAaspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1195 GCGGGGCTGAACCGAGACCGTCCCGTCTGCGCGCGCGGTGGCCAGGTGCTG 1251
RESULT 12
US-11-136-527-2278
; Sequence 2278, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2278
; LENGTH: 2343
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2278
Alignment Scores:
Pred. No.: 1.08e-74 Length: 2343
Score: 781.00 Matches: 162
Percent Similarity: 57.3% Conservative: 62
Best Local Similarity: 41.4% Mismatches: 165
Query Match: 38.2% Indels: 2
DB: 12 Gaps: 2
US-10-673-786A-2 (1-396) x US-11-136-527-2278 (1-2343)
Qy 4 AsnIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPheArgAlaAasp 23
Db 192 CATGTTGAAATGGGACCTCCAGATCCCATCTCTGGAGTACCGAGGCTTCAGWAGAGAT 251
Qy 24 GluAargProGlyLysIleAenLeuGlyIleGlyValTyrLysAaspGluThrGlyLysThr 43
Db 252 ACCAACAGCAGAAAGATGAACCTGGGAGTTGGTGCCTACCGGACGATTAACGGAAGCCT 311
Qy 44 ProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGluThrThrLys 63
Db 312 TACGTGCTCCCAAGTGTTCGAAAGGAGAGGCCAGATTCGTGGGAAAAATTTGGWCAAA 371
Qy 64 AsnTyrLeuGlyIleAaspGlyIleProGluPheGlyArgCysThrGlnGluLeuPhe 83
Db 372 GAATACCTACCATCGGGAGCTGGCTGATTTTGTAAAGCTCTCGAGAACTGGCCCTG 431
Qy 84 GlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGlnThrProGly 103
Db 432 GCGGAGAACCGGAGGTGTTGAAAAAGCGCGGTTTGTAACTGTGACAGACCATTTCCGG 491
Qy 104 ThrGlyAlaLeuAargValAlaAaspPheLeuAlaLysAenThrSerValLysArg--- 122
Db 492 ACTGAGCCCTTGAGGTCGAGCCAGCTTTCTGCAAAAGATTTTTTAAGTTCAGCCGAGAT 551
Qy 123 ValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAlaGlyLeu 142
Db 552 GTCTTTCTGCCAACCATCTCGGGAAACACACGCGCCATCTTCAGGGATCGCGGATG 611
Qy 143 GluValAargGluTyrAlaTyrTyrAaspAlaGluAenHisThrLeuAaspPheAaspAlaLeu 162
Db 612 CAGCTGCAAGGTTATCGATCTACTATGACCCCAAGACTTTCGCGCTTTGACTTCTCTGGAGCC 671
Qy 163 IleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCysCysHis 182

```

```

Db      672  TTAGAAGACATATAAAAAATCCAGAGCAGAGTGTCTCTCTTGTGACGCGCTGCAC 731
Qy      183  AsnProThrGlyLeuAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeuSerVal 202
Db      732  AACCCACGGCGGTGAGCCCGCTCCAGAGCAGTGGAGGAATGGCGCGGTGTGAG 791
Qy      203  GluLysGlyTrpLeuProLeuPheAaspPheAlaTrpGlnGlyPheAlaArgGly---Leu 221
Db      792  AAAAGAATCTCTTCGATCTTTGACATGGCTACCAAGGCTTTGCCAGCGCGCATGT 851
Qy      222  GluGluAaspAlaGluGlyLeuAlaPheAlaAlaMetHisLysGlnLeuLeuValAla 241
Db      852  GATAAGGACGCTGGCGCGCTTCAGGCACTTCATCGAGCAGGCAATCAATGCTGCTG 911
Qy      242  SerSerTySerLysAsnPheGlyLeuTyAsnGluAtrGValGlyValaCysThrLeuVal 261
Db      912  CAATCTATGCCAAGAACATGGCCCTGTACGGTGAGCGGTGGAGCCTTCACTGTGGS 971
Qy      262  AlaAlaAaspSerGluThrValAaspAlaPheSerGlnMetLysAlaAlaAlaAla 281
Db      972  TCAGAGATGCAGAGAACCAAGGGTGGAGTCACAGCTGAAGATCTGTATCGCCCC 1031
Qy      282  AsnTySerAsnProProAlaHisGlyAlaSerValAlaThrIleLeuSerAsnAasp 301
Db      1032  TTGTATTCAACCGCCTCTCAATGAGCGCGGATCGCGCAACCATYCTGAYKCTSSA 1091
Qy      302  AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 321
Db      1092  GACKSCGGAAGCAATGGTTGAGAGAGTGGAAAGCGATGGCTGACCGCATCATCAG 1151
Qy      322  ArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAaspPheIle 341
Db      1152  AGACCCAGTGTCTCCAACTGGAAGAGAGGCTGCTCCCACTGCGGAGCACATC 1211
Qy      342  IleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 361
Db      1212  ACCGACCAAGATCGCATGTTCTGTTCCCGGCTTAAAGCCTGAGCAGGTGGAGCG 1271
Qy      362  ArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 381
Db      1272  ACCAAGAGKCTYAGKCTACATGACAAAGAGTGGTGGATCTCTGKGGCGGGGTACC 1331
Qy      382  ProAspAsnMetAlaProLeuCysGluAlaIle 392
Db      1332  TCTGCAATGTGGGTAMCTGGCCCAAGCCATT 1364

```

RESULT 13

```

; Sequence 82350, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHITHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHKART, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82350
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82350

Alignment Scores:
Pred. No.: 3,43e-69 Length: 1251
Score: 727.50 Matches: 153
Percent Similarity: 57.5% Conservative: 74

```

```

Best Local Similarity: 38.7% Mismatches: 157
Query Match: 35.6% Indels: 11
DB: 7 Gaps: 6
US-10-673-786A-2 (1-396) x US-10-932-182A-82350 (1-1251)
Qy      1  MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db      55  CTGTTCAATACATCGAATTGCTGCCCTGATGCTCTTTTGGTATTAAAGCAAAAGGTAC 114
Qy      21  ArgAlaAaspGluAtrGProGlyLysIleAsnLeuGlyIleGlyValTyLysAaspGluThr 40
Db      115  GGGCAAGATCAACAGTCTACCAAGTGCATCGGGTATCGGGGCTTACAGACGACCAAC 174
Qy      41  GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAenGlu 60
Db      175  GGTAAACCATGGTCTTCCCAAGTGTAAAGCCGCGGAAAG---CTAATTCATAACGAC 231
Qy      61  ThrThrLysAsn-----TyLeuGlyIleAaspGlyIleProGluPheGlyArgCysThr 78
Db      232  AGCTCTCAACACCATGAATACCTCGGTATTACCGGTCTGCCAAGTTTGACATCTAAACGCC 291
Qy      79  GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAaspLysArgAlaAtrGThrAla 98
Db      292  GCCAAGATCATCTTCGGTACCGAATCCGATCGCTTGCAGGAAGACAGAGTAATCTCAGTA 351
Qy      99  GlnThrProGlyGlyThrGlyValAlaLeuArgValAlaAlaAaspPheLeuAlaLysAenThr 118
Db      352  CAATCAGTCTGGTACGGGTGCTCTTCATATATTGGGAAGTTTTTTCAAATTCCTTC 411
Qy      119  SerValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsn 138
Db      412  CCAGATAAATCGTCTATTGCTAAGCTTACTTGGGCAACACACATGCGCCATTTTGTAG 471
Qy      139  SerAlaGlyLeuGluValArgGluTyAlaTyTrpAspAlaGluAsnHisThrLeuAasp 158
Db      472  AATCAAGCTTGAAGAACCGGCACTTACCTTTACTGGGCCAACGAAATAAGCTTTTGGAC 531
Qy      159  PheAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAaspValValLeuPheHis 178
Db      532  CTAAACGGCTTCTTAATGCTATTCAAAGAGCTCCAGAGGGCTCCATTTTCTGTCGAC 591
Qy      179  GlyCysCysHisAsnProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAla 198
Db      592  TCTTGGCGCCATAACCAACTGGTCTGGACCTACTAGTGAACAATGGTTCAAATCGTT 651
Qy      199  GlnLeuSerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyGlnGlyPheAla 218
Db      652  GATGCTATCGCTCAAAAAACCAATCGCTTATTGTACACGCGCTTACCAAGGTTTGGC 711
Qy      219  ArgGly---LeuGluGluAaspAlaGluGlyLeuArgAlaPheAla-----AlaMet 234
Db      712  ACTGGAGATTGGACAAGGATGCTATGCTGCGCTCTAGGTGTGGAGAGCTTTCGACG 771
Qy      235  HisLysGluLeuIleValAlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArg 254
Db      772  GTCCTCCCGCTTGTGTCTGTCAGTCTTTCGCAAGAACGCGCGGTATGTACGGTGGCGT 831
Qy      255  ValGlyAlaCysThrLeuVal-----AlaAlaAaspSerGluThrValAaspArgAla 271
Db      832  GTAGGTTGTTTCCATCTAGCACTTCAAAAAACAAAGCTCAAAAAACAAATATAAAGCCTGCT 891
Qy      272  Phe---SerGlnMetLysAlaAlaIleArgAlaAenTySerAsnProProAlaHisGly 290
Db      892  GTTACATCTCAATTGGCCAAAATCATCTGATGAAGTGTCCCAACCCCGCGCTTACGCG 951
Qy      291  AlaSerValAlaAlaThrIleLeuSerAsnAaspAlaLeuArgAlaIleTrpGluGlnGlu 310
Db      952  GCTAAGATTGCTGCTAAACTGTTGAAACGCCAGATTAACGGAACAGTGGCACAAGAT 1011
Qy      311  LeuThrAspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValAsnThrLeuGln 330
Db      1012  ATGTTTACCATGTCCTCCAGAATTACGAAAATGAGCGCAGCATTAAGAGACCATTTAGTC 1071

```

```
Qy 331 GluLysGlyAlaAsnArgAspPheSerPheIleLeuLysGlnAsnGlyMetPheSerPhe 350
Db 1072 AAGTTGGGCACTCCTGGCACTGGGATCATATAGTAAATCAATCGGGATGTTCTCCTTT 1131
Qy 351 SerGlyLeuThrLysGluGlnValLeuArgLeuArgGluGluPheGlyValTyrAlaVal 370
Db 1132 ACAGGATTGACTCCTCAATGGTTAAACGACTTGAAGAAACCCACGCGAGTTTACTTGTT 1191
Qy 371 AlaSerGlyArgValAsnValAlaGlyMetThrProAspAsnMet 385
Db 1192 GCCTCAGGTAGAGCTTCTATTGCTGGATTGAATCAAGGAAACGTTG 1236

RESULT 14
US-10-932-182A-82350
; Sequence 82350, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 82350
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82350

Alignment Scores:
Pred. No.: 3.43e-69 Length: 1251
Score: 727.50 Matches: 153
Percent Similarity: 57.5% Conservative: 74
Best Local Similarity: 38.7% Mismatches: 157
Query Match: 35.6% Indels: 11
Db: 7 Gaps: 6

US-10-673-786A-2 (1-396) x US-10-932-182A-82350 (1-1251)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 55 CTGTTCAATAACATCGAATTGCTGCCCTCCCTGATGCCCTTTTGGTATTAAAGCAAGGTAC 114
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 115 GGGCAAGATCAACGTCGTACCAAGTCGACTTGGGTATCGGGCTACAGAGACGACAAC 174
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 175 GGTAAACCATCGGTCTTCCCAAGTGTAAAGCCGCGAAAG--CTAATTCAATAACGAC 231
Qy 61 ThrThrLysAsn-----TyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThr 78
Db 232 AGCTCCTCAACCATCAATGAAATACCTCGTATTACCGGTCTGGCAAGTTTGACATCTAACGCC 291
Qy 79 GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAla 98
Db 292 GCCAAGATCATCTTCGGTACCAATCCGATCCCTTGCAGGAAGACAGAGTAATCTCAGTA 351
Qy 99 GlnThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThr 118
Db 352 CAATCACTGTCTGGTACGGGTCTCTCATATATTTCGGAAGTTTTCATAAATCTTTC 411
Qy 119 SerValLysArgValTyrValSerAsnProSerThrProAsnHisLysSerValPheAsn 138
Db 412 CCAGATAAATCGGTCTATTGTTGAAGCTTACTTGGGCGCAACCATCGGCCATTTTGGAG 471
```

```
Qy 139 SerAlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAsp 158
Db 472 AATCAAGGCTTGAACACGGCGACTTACCCTTACTTGGGCCAACGAACCTTAAGTCTTTGGAC 531
Qy 159 PheAspAlaLeuIleAsnSerLeuAsnGlnAlaGlyAspValValLeuPheHis 178
Db 532 CTAAACGGCTTTTAAATGCTATTCAAAAAGCTCCAGAGGCTCCATTTTCGTTCTGCAC 591
Qy 179 GlyCysCysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAla 198
Db 592 TCTTGGCCCATAAACCAACTGGTCTGACCCCTACTAGTGAACAATGGTTCAAATCGTT 651
Qy 199 GlnLeuSerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnPheAla 218
Db 652 GATGCTATCGCCTCAAAAAACCAACATCGCTTATTTCACACCCGCTACCAAGGTTTGGCC 711
Qy 219 ArgGly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAla-----AlaMet 234
Db 712 ACTGGAGATTGGACAAAGGATGCTATGCTGTGGCTCTAGGTGGAGAACGTTTCGACG 771
Qy 235 HisLysGluLeuIleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArg 254
Db 772 GTCTCTCCGCTCTTGTCTGTCTGTCAGTCTTTGGCCAAAGAACGCCGCTATGTACGGTGAGCGT 831
Qy 255 ValGlyAlaCysThrLeuVal-----AlaAlaAspSerGluThrValAspArgAla 271
Db 832 GTAGGTTGTTTCCATCTAGCATTACAAAAACAAGCTCAAAACAACAACTATATAAGCCCTGCT 891
Qy 272 Phe---SerGlnMetLysAlaAlaIleArgAlaAsnTyrSerAsnProAlaHisGly 290
Db 892 GTTACATCTCAATTGGCCAAAATCATTCGTAGTGAAGTGTCCAAACCCACCGCCCTACGGC 951
Qy 291 AlaSerValValAlaThrIleLeuSerAsnAspAlaLeuArgAlaIleTyrGluGlnGlu 310
Db 952 GCTAAGATTGCTCGTAACTGTTGGAAACGCCAGAAATTAACGGAAACAGTGCACAAGAT 1011
Qy 311 LeuThrAspMetArgGlnArgIleGlnArgMetArgGlnLeuPheValAsnThrLeuGln 330
Db 1012 ATGGTTACCATGTCTCTCAGAAATTTACGAAATTAGGCACCGCATTTAAGAGACCATTTAGTC 1071
Qy 331 GluLysGlyAlaAsnArgAspPheSerPheIleIleLysGlnAsnGlyMetPheSerPhe 350
Db 1072 AAGTTGGGCACTCCTGGCACTGGGATCATATAGTAAATCAATCGGGGATGTTCTCCTTT 1131
Qy 351 SerGlyLeuThrLysGluGlnValLeuArgLeuArgGluGluPheGlyValTyrAlaVal 370
Db 1132 ACAGGATTGACTCCTCAATGGTTTAAACGACTTGAAGAAACCCACGCGAGTTTACTTGTT 1191
Qy 371 AlaSerGlyArgValAsnValAlaGlyMetThrProAspAsnMet 385
Db 1192 GCCTCAGGTAGAGCTTCTATTGCTGGATTGAATCAAGGAAACGTTG 1236

RESULT 15
US-10-932-182A-4671
; Sequence 4671, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 4671
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-4671
```


Qy	1	MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe	20
Db	145	CTATTCAACAACATCGAACTGCTCCCGCCTGACGCCCTTTTGGCATCAAGCAAAAGATAC	204
Qy	21	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr	40
Db	205	GGTCAAGACCAACGACCAACCAAGGTGCACTGGGTATCGGAGCCTACAGAGACAATAAC	264
Qy	41	GlyLysThrProValIleThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu	60
Db	265	GGTAAGCCATGGGTCCTGCCAAGTGTCAAGCCGCGGAAGAGCTTATCCATAGCAGCA	324
Qy	61	Thr---ThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln	79
Db	325	TCCTACAATCAAGAGTACTCTCGGATCACTGGCTGCCTCTTTGACGTCCAACGCCGCC	384
Qy	80	GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln	99
Db	385	AAGATCATCTTCGGTACGCAATCCGCCCTTGCAGGAGGACAGAGTAGTAATTTCCGTG	444
Qy	100	ThrProGlyThrGlyValAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer	119
Db	445	TCGCTGTCGGTACTGGAGCCCTTCATATAGCCGCAAGTTCTTCTCAAGTTCTATCCA	504
Qy	120	ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer	139
Db	505	GACAACTGGTCTACCTGTCCAAGCCACCTGGGCAAAACCACACGGCCATCTTTGAA	564
Qy	140	AlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPhe	159
Db	565	CAAGGTTTGAATAACACCAACCTACCCCTACTGGGCCAATGAGACCAAGTCTTGATTA	624
Qy	160	AspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly	179
Db	625	GACGGCTTCCTCGCGCCGCTCCAAAGCCGCCAGAGGTTCCGTCTTTGTTCTTCACTCT	684
Qy	180	CysCysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrpGlnThrLeuAlaGln	199
Db	685	TGTCTCACAACTTACTGGTCTGGACCCCACTAACGACCAATGGCCGAGATTTCGCAG	744
Qy	200	LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg	219
Db	745	GCCATTGCCGCCAAGACCAAGTTCGCCCTATTTCGACACGGGTACAGGGTTTCGTACA	804
Qy	220	Gly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAla-----AlaMethHis	235
Db	805	GGTCACCTGGACAAAGACGCTACGCCGTGGTCTGGGTGGAGAGAACTATCCACAGTC	864
Qy	236	LysGluLeuIleValAlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgVal	255
Db	865	TCTCCCGTCTTTGTCTGTCACGTCTGTTCCCAAGAACGCTGGTATGTACGGTGAGCGTG	924
Qy	256	GlyValCysThrLeuVal-----AlaAlaAspSerGluThrValAspArgAlaPhe	272
Db	925	GGTTGTTTCCATCTAGCGCTCACAAAGCCCAAAACAGACCATCAAGCCCCCTGTC	984
Qy	273	---SerGlnMetLysAlaAlaIleArgAlaAsnTyrSerAsnProAlaHisGlyAla	291
Db	985	ACCTCCCAATTGGCCAAGATCATACGTAGTGAAGTGTCCATCCACCGCCCTACGGTGC	1044
Qy	292	SerValValAlaThrIleLeuSerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeu	311

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:34:18 ; Search time 6793 Seconds
(without alignments)
3223.239 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITRAPADPILGLADLF.....VAGMTPDNNAFLCAIVAVL 396

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -US10673786/runat_14032006_135615_22715/app_query.fasta_1
-Q=/abss/ABSSWEB_spool/US10673786/runat_14032006_135615_22715/app_query.fasta_1 -LOOFCPL=0
-DB=Pending Patents NA Main -QFMT=fastap -SUFFIX=brnmp -MINMATCH=0.1 -LOOFCPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss04 -USER=US10673786@CGN 1 1 7336 @runat_14032006_135615_22715
-NCPU=6 -ICPU=3 -NO MAP -NEG SCORE=0 -WAIT -DSBLOK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA Main:
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	1191	50	US-10-369-493-24472
2	2045	100.0	1191	54	US-10-673-786A-1
3	2045	100.0	1191	78	US-60-360-039-24472
4	2045	100.0	1191	82	US-60-638-099-7
5	2045	100.0	14759	33	US-09-801-563-87
6	2045	100.0	14759	62	US-10-893-671-87
7	2045	100.0	4639675	83	US-60-709-960-1
8	2039	99.7	1576	82	US-60-643-717-112
					Sequence 24472, A
					Sequence 1, Appli
					Sequence 24472, A
					Sequence 7, Appli
					Sequence 87, Appli
					Sequence 1, Appli
					Sequence 112, App

9	2011	98.3	1191	70	US-11-114-922-75	Sequence 75, Appl	1	MetPheGluAsnIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe	20
10	1931	94.4	1209	51	US-10-446-203-4902	Sequence 4902, Ap	1	ATGTTTGAACAATACCGCGCTCTCTGCGACCCGATTTCTGGCGCTGGCGATCTGTTT	60
11	1930	94.4	1209	22	US-09-252-691-1606	Sequence 1606, Ap	21	ArgAlaaspGluArgProGlyLysIleAsnLeuGlyLeuValTyrLysAspGluThr	40
12	1930	94.4	1209	22	US-09-252-691C-1606	Sequence 1606, Ap	61	CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	120
13	1930	94.4	1209	51	US-10-417-886-1606	Sequence 1606, Ap	41	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu	60
14	1930	94.4	1218	22	US-09-252-691-1616	Sequence 1616, Ap	121	GGCAAAACCCGCTACTGACGCGTGAAGGCTGAACAGTATCTGCTCGAAATGAA	180
15	1930	94.4	1218	22	US-09-252-691C-1616	Sequence 1616, Ap	61	ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu	80
16	1930	94.4	1218	51	US-10-417-886-1616	Sequence 1616, Ap	181	ACCACCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	240
17	1654	80.9	1188	50	US-10-369-493-44812	Sequence 44812, A	81	LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr	100
18	1654	80.9	1188	78	US-10-369-493-44812	Sequence 44812, A	241	CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGACACT	300
19	1654	80.9	1191	71	US-11-150-804-4967	Sequence 4967, Ap	101	ProGlyGlyThrGlyAlaLeuArgValAlaAlaAaspPheLeuAlaLysAsnThrSerVal	120
20	1654	80.9	1191	81	US-60-579-902-4967	Sequence 4967, Ap	301	CCGGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATACACGCGTT	360
21	1654	80.9	1191	81	US-60-579-902A-4967	Sequence 4967, Ap	121	LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla	140
22	1654	80.9	524707	71	US-11-150-804-14985	Sequence 14985, A	361	AAGCGTGTGGTGGTGACCAACCCCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA	420
23	1654	80.9	524707	81	US-60-579-902-14985	Sequence 14985, A	141	GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp	160
24	1654	80.9	524707	81	US-60-579-902A-14985	Sequence 14985, A	421	GGTCTGAAGTTCGTGAATACGCTTATTATGATCGGAAATACACCTCTTGACTTCGAT	480
25	1650	80.7	1191	72	US-11-216-782-3916	Sequence 3916, Ap	161	AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys	180
26	1650	80.7	1191	82	US-60-606-098-3916	Sequence 3916, Ap	481	GCATGATTAAACGCTGAATGAAGCTCAGCTGGCGACAGTAGTGCTGTTCCATGGCTGC	540
27	1650	80.7	171354	72	US-11-216-782-16894	Sequence 16894, A	181	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
28	1650	80.7	171354	82	US-60-606-098-16894	Sequence 16894, A	541	TGCCATAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAACACACATGGCAACATC	600
29	1343	65.7	1191	82	US-60-691-214-2033	Sequence 2033, Ap	201	SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
30	1343	65.7	7977	2	PCT-US03-17092-536	Sequence 536, App	601	TCCGTTTGAGAAAGCGTGGTTACCGCTGTTGACTTCGCTTACCGAGGTTTGGCCGCGGT	660
31	1343	65.7	7977	61	US-10-795-159-536	Sequence 536, App	221	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal	240
32	1343	65.7	7977	79	US-60-453-134-536	Sequence 536, App	661	CTGGAAGAAGATGCTGAAGGACTGGCGGCTTTCGCGCTATGCATAAAGAGCTGATTGTT	720
33	1343	65.7	7977	82	US-60-691-214-536	Sequence 536, App	241	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
34	1343	65.7	908766	2	PCT-US04-07001-536	Sequence 536, App	721	GCCAGTTTCTACTCTAAAAACTTTGGCTGTGTACAAACGAGCGTGTGGCGCTTGACTCTG	780
35	1343	65.7	908766	61	US-10-795-159-685	Sequence 685, App	261	ValAlaAlaaspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
36	1343	65.7	908766	82	US-60-691-214-685	Sequence 685, App	781	GTTGCTCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGGATTCGC	840
37	1343	65.7	908766	82	US-60-691-214-772	Sequence 772, App	281	AlaAsnTyrSerAsnProAlaHisGlyValaSerValValAlaThrIleLeuSerAsn	300
38	1343	65.7	1913428	82	US-08-426-787-1	Sequence 1, Appl	841	GCTAACTACTCTAACCCACGACGACGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	900
39	1325	64.8	1830121	14	US-10-329-670-1	Sequence 1, Appl	301	AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg	320
40	1325	64.8	1830121	49	US-10-329-670-1	Sequence 1, Appl	901	GATCGGTTTACGTGCGATTTGGGAACAAGAGCTGACTGATGATGCGCCGCGTATTTCAGCGT	960
41	1325	64.8	1830121	64	US-10-981-687-1	Sequence 23918, A	321	MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe	340
42	1318	64.4	966	50	US-10-369-493-23918	Sequence 23918, A	961	ATGCGTCAGTGTGTTGTCGTCATACGCTCAGGAAAGCGCGCAAAACCGCGACTTCAGCTTT	1020
43	1318	64.4	966	78	US-60-360-039-23918	Sequence 23918, A	341	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
44	1318	64.4	4404	34	US-09-897-516-374	Sequence 1179, Ap	1021	ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGCACAAAAGAACACAGTGTCTGCGT	1080
45	1318	64.4	4404	34	US-09-897-516-1179	Sequence 1179, Ap			

ALIGNMENTS

RESULT 1
US-10-369-493-24472
; Sequence 24472, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24472
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24472

Alignment Scores:
Pred. No.: 1.4e-314 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 50 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-369-493-24472 (1-1191)

QY 361 LeuArgGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCAAGAGTTTGGCGTATATATGCGTTCTGTCGCTAAATGTCGCGGATG 1140
 QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaValAlaValLeu 396
 Db 1141 ACACCAAGATACATGCTCCGCTGTGCGAAGCATTTGTGGCAGTGTCTG 1188

RESULT 2

US-10-673-786A-1
 ; Sequence 1, Application US/10673786A
 ; GENERAL INFORMATION:
 ; APPLICANT: AGHVERDIAN, VALERY ZAVENOVICH
 ; APPLICANT: SAVRASOVA, EKATERINA ALEKSEEVNA
 ; APPLICANT: KAPLAN, ALLA MARKOVNA
 ; APPLICANT: LOBANOV, ANDREY OLEGOVICH
 ; APPLICANT: KOZLOV, YURI IVANOVICH
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-THREONINE USING BACTERIA
 ; TITLE OF INVENTION: BELONGING TO THE GENUS ESCHERICHIA
 ; FILE REFERENCE: US-115
 ; CURRENT APPLICATION NUMBER: US/10/673,786A
 ; PRIORITY FILING DATE: 2003-09-30
 ; PRIOR APPLICATION NUMBER: PCT/JP03/02067
 ; PRIOR FILING DATE: 2003-02-27
 ; PRIOR APPLICATION NUMBER: 2002104983
 ; PRIOR FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1191
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1191)
 US-10-673-786A-1

Alignment Scores:

Pred. No.: 1,4e-314 Length: 1191
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 54 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-673-786A-1 (1-1191)

QY 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 1 ATGTTTGAGAACATTACCGCCGCTCTGCGGACCCGATTTCTGGGCTGGCCGATCTGTT 60
 QY 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 61 CGTGGCGATGAACGTCGCGCAAAATTAACCTCGGGATTTGTGTCTATAAAGATGAGACG 120
 QY 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
 Db 121 GGCAAAACCCGGTACTACCGCGGTGAAAAGGTGAACAGTATCTGCTCGAAATGAA 180
 QY 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 181 ACCACCAAAATTAACCTCGGCATTGACGCGCATCCCTGAATTTGGTGGTGCACACTCAGAA 240
 QY 81 LeuLeuPheGlyLysGlySerIleLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGCACAGAT 300
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 301 CCGGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTGGCAAAATACACGCGTT 360
 QY 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140

Db 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCGAACCATATAAGAGCGCTCTTTAACTCTGCA 420
 QY 141 GlyLeuGluValArgGlyTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 421 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATACACACTCTTGACTTCGAT 480
 QY 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GCACCTGATTAAACAGCTCAATGAAGCTCAGGCTGCGAGCTAGTGTGTTCATGCTGCTGC 540
 QY 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
 Db 541 TGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACAATGCAACACTTGGCACAATC 600
 QY 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 601 TCCGTTGAGAAGGCTGGTTACCGCTTTTGACTTCGCTTACAGGTTTTCGCCGCTGGT 660
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuVal 240
 Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGCGCTATGCATAAAGAGCTGATTGT 720
 QY 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGTTCTTACTCTAAAACCTTTGCGCTGTACACAGCGGTGTGGCGCTTTGTACTCTG 780
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGCTCGGACAGTGAACCGTTGATCGGATTCAGCAATTAAGCAAGCGCGGATTCGC 840
 QY 281 AlaAsnTyrSerAsnProAlaHisGlyLysValValAlaThrIleLeuSerAsn 300
 Db 841 GCTAACTACTCTAACCCACAGCACACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
 QY 301 AspAlaLeuArgAlaIleTyrGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 901 GATCGTTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATGCGCAGCGTATTCAGCGT 960
 QY 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAlaAsnArgAspPheSerPhe 340
 Db 961 ATGCGTCAGTTGTTCTGTAATACGCTGCGGAAAGGCGCAACCCGCGACTTCAGCTTT 1020
 QY 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATCATCAACAGACGCGCATGTTCTCTCTCAGTGGCTGACAAAAGAACAGTGTGCGT 1080
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGTAAATGTGCGCGGATG 1140
 QY 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACCAATAACATGGCTCCGCTGTCGGAAGCATTTGTGGCAGTGTCTG 1188

RESULT 3

US-60-360-039-24472
 ; Sequence 24472, Application US/60360039
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Chen, Xianfeng
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)A
 ; CURRENT APPLICATION NUMBER: US/60/360,039
 ; CURRENT FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 24472
 ; LENGTH: 1191
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli

US-60-360-039-24472

Alignment Scores:

Pred. No.: 1.4e-314 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 78 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-360-039-24472 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTGCGACCGGATTCGGCGCTGGCGGATCTGTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyLeuGlyValTyrLysAspGluThr 40
Db 61 CGTCCGATGAACGTCGCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAAACCCCGGTACTGACCGGCTGAAAGGCTGAAAGGCTGACAGTATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenTyrLeuGlyLysIleAspGlyLeuProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTAACCTCGCATTCGCGCATCCCTGGAATTTGGTGGCTGCACCTCAGGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAGGTAGCGCCCTGATCAATGACAAACGCTGTCGACCGCACAGACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGCGACTGGCGGCTAGCGTGGCTGCCGATTTCTGGGCAAAATTAACCGGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 361 AAGCGTGTGGTGAGCAACCCCAAGCTGCGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTTCTGGAATACGCTATTATGATCGGAAATCACTCTTTGACTTCGAT 480
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCATGTATTACAGCCTGAATGAGCTCAGGCTGGCGACGTAGTGTCTTCATGGCTGC 540
Qy 181 CysHisAenProThrGlyLysAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCCCAACCGGTATCGACCTACGCTGGAACAATGGCAACACTGGCACAACTC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCCGTTGAGAAGGCTGGTTTACCGCTGTTTACCTTCGCTTACCAAGGGTTTTGCCCGCTGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysLysGluLeuIleVal 240
Db 661 CTGAAGAAGATGCTGAAGAGCTGCGCGCTTTCGCGGCTATGCAATAAAGAGCTGATTGTT 720
Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTCTCTACTCTAANAACCTTTGGCCTGTACACGAGCGGTGTGGCGCTGTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGTCGCCGAGTGAACCGCTTGTATCGCGCATTCAGCCCAATGAAGCGCGGATTCGC 840
Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 GCTAACTACTCTAACCACACAGCACCGCGCTTCTGTTGTGGCACCATCTCTAGCNAAC 900
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320

Db 901 GATGCGTTACGTGCGATTTCGGGAACAAGAGCTGACTGATATGCCCAAGCTATTTCAGCGT 960
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnGluLysGlyAlaAenArgAspPheSerPhe 340
Db 961 ATGCGTCAGTTGTTGCTCAATACGCTGACGAGAAAGGCGCAAAACCGGACTTCAGGCTTT 1020
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGACAAAGAACAGTCTGCGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTCGTCGCTGAAATGTGGCGGATG 1140
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACCAATACATGGCTCCGCTGTGCAAGCGATTCTGGCAGTGCTG 1188

RESULT 4

US-60-638-099-7

; Sequence 7, Application US/60638099

; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D

; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes

; FILE REFERENCE: 38-21(53720)

; CURRENT APPLICATION NUMBER: US/60/638,099

; CURRENT FILING DATE: 2004-12-21

; NUMBER OF SEQ ID NOS: 48056

; SEQ ID NO 7

; LENGTH: 1191

; TYPE: DNA

; ORGANISM: Escherichia coli

US-60-638-099-7

Alignment Scores:

Pred. No.: 1.4e-314 Length: 1191
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 82 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-638-099-7 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTGCGACCGGATTCGGCGCTGGCGGATCTGTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyLysValTyrLysAspGluThr 40
Db 61 CGTCCGATGAACGTCGCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 121 GGCAAAACCCCGGTACTGACCGGCTGAAAGGCTGAAAGGCTGACAGTATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenTyrLeuGlyLysIleAspGlyLeuProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTAACCTCGCATTCGCGCATCCCTGGAATTTGGTGGCTGCACCTCAGGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGTCGACCGCACAGACT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGCGACTGGCGGCTAGCGCTGGCTGCCGATTTCTGGGCAAAATTAACCGGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAsnSerAla 140
Db 361 AAGCGTGTGGTGAGCAACCCCAAGCTGCGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420

QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAAsnHisThrLeuAspPheAsp 160
 Db 421 GGTCTGGAAGTTTCGTAATACGCTTATTATGATGCGAAATTCACACACTTGTGACTTCGAT 480
 QY 161 AlaLeuIleAAsnSerLeuAAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 481 GCACGTGATTAACAGCGCTGAATGAAGCTCAGGCTGCGACGTAGTCTGTTCATCGGCTGC 540
 QY 181 CysHisAAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 541 TCCCATTAACCCCAACCGGTATGACCTTACGCTGCGAACAATGCAACACTGGCACACTC 600
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlnPheAlaArgGly 220
 Db 601 TCCGTTTGAGAAGGCTGTTTACCGCTGTTTGAATTCGCTTACACGGGTTTGGCCGTTGGT 660
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisGlyGluLeuIleVal 240
 Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGGCTATGCATAAAGAGAGCTGATTGTT 720
 QY 241 AlaSerSerTyrSerLysAAsnPheGlyLeuTyrAAsnGluArgValGlyAlaCysThrLeu 260
 Db 721 GCCAGTCTCTACTCTAAACAACTTGGCCTGTACAACGAGCGGTGGCGCTTGTACTCTG 780
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 781 GTTGTGCGGACAGTGAACCGTTGATCGCGATTCAGCCAAATGAAGCGGCGATTGCG 840
 QY 281 AlaAAsnTyrSerAAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAAsn 300
 Db 841 GCTAACTACTCTAACCCACGACACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAAC 900
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 901 GATGCGTTTACGCGCATTTGGGAACAAGAGCTGACTGATATGCGCCGCGATTTACGCGT 960
 QY 321 MetArgGlnLeuPheValAAsnThrLeuGlnGluLysGlyAlaAAsnArgAspPheSerPhe 340
 Db 961 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAGGCGGCAACCGCGACTTTCAGCTTT 1020
 QY 341 IleIleLysGlnAAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1021 ATCATCAACAGACAGCGCATGTTCTCTTCACTGCGCTGACAAAAGAACAGTGTGCGT 1080
 QY 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAAsnValAlaGlyMet 380
 Db 1081 CTGCGCGAAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1140
 QY 381 ThrProAspAAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1141 ACACCATTAACATGGCTCCGCTGCGAAGGCAATGTGGCAGTCTG 1188

RESULT 5

US-09-801-563-87/c
 ; Sequence 87, Application US/09801563
 ; GENERAL INFORMATION:
 ; APPLICANT: Levy, Stuart, et. al.
 ; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
 ; FILE REFERENCE: PRZ-043
 ; CURRENT APPLICATION NUMBER: US/09/801,563
 ; PRIORITY FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/188,362
 ; PRIOR FILING DATE: 2000-03-10
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 87
 ; LENGTH: 14759
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-801-563-87

Alignment Scores: 2.04e-312 Length: 14759
 Pred. No.:

Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservativity: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 Ds: 33 Gaps: 0
 US-10-673-786A-2 (1-396) x US-09-801-563-87 (1-14759)
 QY 1 MetPheGluAAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 2755 ATGTTTGGAACATTAACCGCGCTCTCGCCGACCGATCTGGGCTGGCGATCTGTTT 2696
 QY 21 ArgAlaAspGluArgProGlyLysIleAAsnLeuGlyIleGlyValTyrLysAspGluThr 40
 Db 2695 CGTCCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 2636
 QY 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAAsnGlu 60
 Db 2635 GGCAAAACCCCGGTACTGACCGAGCTGAAAAGGCTGAACAGTATCTCTCGAAAATGAA 2576
 QY 61 ThrThrLysAAsnTyrLeuGlyIleAspGlyIleProGluPheGlyValArgCysThrGlnGlu 80
 Db 2575 ACCACCAAAATTAACCTCGGCATTTGACGCGATCCCTGAATTTGGTCGCTGCACCTCAGNA 2516
 QY 81 LeuLeuPheGlyLysGlySerAlaLeuIleAAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 2515 CTGCTGTTTGTAAAGGTAGCGCTGATCAATGACAAACGCTGTCGACGCGCACAGACT 2456
 QY 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAAsnThrSerVal 120
 Db 2455 CCGGGGGGCACTGCGCGACTACGCGTGGCTCCGATTTCTTGGCAAAAAATACACAGCGTT 2396
 QY 121 LysArgValTrpValSerAAsnProSerTrpProAAsnHisLysSerValPheAAsnSerAla 140
 Db 2395 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCAATAAGAGCGTCTTTAACTCTGCA 2336
 QY 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAAsnHisThrLeuAspPheAsp 160
 Db 2335 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAATATCACACTTCTGACTTCGAT 2276
 QY 161 AlaLeuIleAAsnSerLeuAAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2275 GCACGTGATTAACAGCGTGAATGAAGCTCAGGCTGCGACGCTAGTGTCTGTTCATGGCTGC 2216
 QY 181 CysHisAAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 2215 TGCCATAACCAACCGGTATCGACCTACGCTGGAACCAATGGCAACACTGGCAACATC 2156
 QY 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
 Db 2155 TCCGTTTGAGAAGGCTGGTTACCGCTGTTTGAATTCGCTTACCAGGCTTTTGGCCGTTGT 2096
 QY 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 2095 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCATAAAGAGAGCTGATTGTT 2036
 QY 241 AlaSerSerTyrSerLysAAsnPheGlyLeuTyrAAsnGluArgValGlyAlaCysThrLeu 260
 Db 2035 GCCAGTCTCTACTCTAAAAACCTTTGGCCTGTACAAACGAGCGGTTTGGCGCTTGTACTCTG 1976
 QY 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 1975 GTTGTGCGGACAGTGAACCGTTGATCGGCATTTACGCCAATGAAGCGCGGATTGCG 1916
 QY 281 AlaAAsnTyrSerAAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAAsn 300
 Db 1915 GCTAACTACTCTAACCCACGACGACGCGCTTCTGTTGTTGGCCACCATCTCTGAGCAAC 1856
 QY 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 1855 GATGCGTTTACGCGATTTGGGAACAAGAGCTGATGATATGCGCGACGCTATTTCAGCGT 1796
 QY 321 MetArgGlnLeuPheValAAsnThrLeuGlnGluLysGlyAlaAAsnArgAspPheSerPhe 340

Db 1795 ATGGCTCAGTTGTTGCTCAATACGCTGCAGGAAAGCGCAACCCGCGACTTCAGCTTT 1736
 Qy 341 lletleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1735 ATCATCAACAGACGCGATGTTCTCCTTCAGTGGCTGACAAAAGAACAGTCTGCGT 1676
 Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1675 CTGGCGAAGAGTTGGCGTATATCGGTTGCTTCTGCTCGCTGAAATGTGGCGGGATG 1616
 Qy 381 ThrProAspAsnMetAlaProLeuGluGluAlaValAlaValLeu 396
 Db 1615 ACACAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1568

RESULT 6

US-10-893-671-87/c
 ; Sequence 87, Application US/10893671
 ; GENERAL INFORMATION:
 ; APPLICANT: Levy, Stuart, et. al.
 ; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
 ; FILE REFERENCE: PK2-043
 ; CURRENT APPLICATION NUMBER: US/10/893,671
 ; CURRENT FILING DATE: 2004-07-15
 ; PRIOR APPLICATION NUMBER: US/09/801,563
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/188,362
 ; PRIOR FILING DATE: 2000-03-10
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 87
 ; TYPE: DNA
 ; LENGTH: 14759
 ; ORGANISM: Escherichia coli
 US-10-893-671-87

Alignment Scores:

Pred. No.: 2,046-312 Length: 14759
 Score: 2045.00 Matches: 396
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 62 Gaps: 0

US-10-673-786a-2 (1-396) x US-10-893-671-87 (1-14759)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 2755 ATGTTTGAGAACATTACCGCGCTCTCGCCAGCCGATTCGGGCTGGCGACTGTTT 2696
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr 40
 Db 2695 CGTGGCGATGAACGTCGCGCAAAATTAACCTCGGATTTGGTGICTATAAAGATGAGACG 2636
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLeuLeuGluAsnGlu 60
 Db 2635 GCGAAACCCCGGTACTACCGAGCTCCGCGCAAAATTAACCTCGGATTTGGTGICTATAAAGATGAGACG 2576
 Qy 61 ThrThrLysAsnTyLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 2575 ACCACCAAAATTAACCTCGGATTTACCGGCTATCGCGCATCCCTGAAATTTGGTGGCTGCACACGAGAA 2516
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
 Db 2515 CTGCTGTTTGTAAAGTAGGCGCTTCATGATGACAAACGCTGCTCGACCGCACAGACT 2456
 Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
 Db 2455 CCGGGGGCGACTGGCGGCATACGCGTGCTGCGGATTTCTGCGCAAAATTAACCGAGCTT 2396
 Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
 Db 2395 AAGCGTGTGGGTGAGCAACCCAGAGCTGGCGCAACCAAGAGCGCTTTTAACTCTGCA 2336

Qy 141 GlyLeuGluValArgGluTyAlaTyTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
 Db 2335 GGTCTGGAAGTTCTGTAATACGCTTATATGATCGCGAAATACACTCTTGACTTCGAT 2276
 Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
 Db 2275 GCACCTGATTAACAGCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 2216
 Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 2215 TGCCATAACCCACCGGTATCGACCTTACCTGGAACAATGGCAAACTACCTGGCACAACTC 2156
 Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAlaArgGly 220
 Db 2155 TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGGCTTCCCTTACCGAGGTTTGGCCGCTGT 2096
 Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
 Db 2095 CTGGAAGAAGATGCTGAGGACTTGGCGCTTTCGGGCTATGCATTAAGAGCTGATTGT 2036
 Qy 241 AlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCysThrLeu 260
 Db 2035 GCCAGTTCTTACTCTAAAACTTTGGCTCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG 1976
 Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
 Db 1975 GTTGCTGCCGACAGTGAACCTTGTATCGCGCATTCAGCCAAATGAAAGCGCGGATTCGC 1916
 Qy 281 AlaAsnTySerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
 Db 1915 GCTAACTACTCTAACCCACCAGCACGCGCTTCTGTTGTGCACCATCTCTGAGCAAC 1856
 Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
 Db 1855 GATGCGTTACGTCGGATTTCGGAACAAGAGCTGACTCATATGCGCCAGCGTATTTCAGCGT 1796
 Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
 Db 1795 ATGCGTCAGTTGTTGCTCAATACGCTCAGGAAAGGCGCAAAACCGCGACTTCAGCTTT 1736
 Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
 Db 1735 ATCATCAACAGACGCGCATGTTCTCTCTCAGTGGCTTGACAAAAGAACAGTCTGCGT 1676
 Qy 361 LeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
 Db 1675 CTGGCGCAAGAGTTTGGCTATATGCGGTTGCTTCTGTCGCTAAATGTGGCGGGATG 1616
 Qy 381 ThrProAspAsnMetAlaProLeuGluGluAlaIleValAlaValLeu 396
 Db 1615 ACACAGATAACATGGCTCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1568

RESULT 7

US-60-709-960-1/c
 ; Sequence 1, Application US/60709960
 ; GENERAL INFORMATION:
 ; APPLICANT: Frederick, Blattner
 ; TITLE OF INVENTION: REDUCED GENOME E. COLI
 ; FILE REFERENCE: 02730,0017.PZUS00
 ; CURRENT APPLICATION NUMBER: US/60/709,960
 ; CURRENT FILING DATE: 2005-08-20
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 1
 ; LENGTH: 4639675
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-60-709-960-1
 Alignment Scores:
 Pred. No.: 1,9e-307 Length: 4639675
 Score: 2045.00 Matches: 396

Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 83 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-709-960-1 (1-4639675)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 984932 ATGTTTGGAGAACATTACCGCGCTCTCTGCCGACCGGATCTGGCGCTGGCGATCTGTT 984873

Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 984872 CGTCCGATGAACGTCGCGCAAAATTAACTCTCGGATTTGTTCTATAAGATGAGACG 984813

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 984812 GGCNAACCCCGGTACTACCGAGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 984753

Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 984752 ACCACAAAATTAATCTCGCATTTGACGCGATCCCTGAAATTTGGTGGCTGCACTCAGAA 984693

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 984692 CTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGTCCTCGCACGCGACAGACT 984633

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 984632 CGCGGGGCGACTGGCGCACTACGCGTGGCTGCGGATTTCTGCGCAAAATACACGCGTT 984573

Qy 121 LysArgValTropValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 984572 AAGCGTGTGGGTGAGCAACCAAGCTGGCGCAACCAATGAGCGTCTTTAACTCTGCA 984513

Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 984512 GGTCTGGAAGTTCGTGATACGCTATTATGATGCGGAAATACACACTCTTGACTTCGAT 984453

Qy 161 AlaLeuLysAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 984452 GCACGTGATTAACAGCGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCCATGGCTGC 984393

Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 984392 TGCCATAACCCCAACCGGTATGACCTACGCTGGCAACATGTCGAACACTGGCAACATC 984333

Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 984332 TCCGTTGAGAAAGCTGTGTACCGCTTTGACTTCGCTTACACGGGTTTTCGCCGCTGGT 984273

Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuVal 240
Db 984272 CTGGAAGAAGATGCTGAAGCACTGGCGCTTTTCGCGCTATGCATAAAGAGCTGATGTT 984213

Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 984212 GCCAGTTCCTACTCTAAAAAATTTGGCGCTGTACAACGAGCGTGTGGCGCTGTACTCTG 984153

Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 984152 GTTGCTGCCGAGTGAACCGTGTGATCGCGATTTACGCCAATGAAGCGCGGATTCGC 984093

Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 984092 GCTAACTACTTAACCCACGACGACGCGCTTCTGTTGTTGCCACCATCTCGAGCAAC 984033

Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 984032 GATCGGTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTACGCGT 983973

Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyAlaAsnArgAspPheSerPhe 340
Db 983973

Db 983972 ATGCGTCAGTTGTTCTGTCATACGCTGAGAAAGGCGCAACCCGCACTTACGCTTT 983913

Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 983912 ATCATCAACAGAACGCGATGTTCTCTTCACTGGCTGACAAAAGAACAGTGTGCGT 983853

Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 983852 CTGCGCAAGAGTTTGGCGTATATCGGTTCTCTGTCGCTGTAATGTGCGCGGATG 983793

Qy 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 983792 ACACAGATAACATGGCTCCGCTGCGAAGCGATTGTGGCAGTCTG 983745

RESULT 8
US-60-643-717-112
; Sequence 112, Application US/60643717
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)A
; CURRENT APPLICATION NUMBER: US/60/643,717
; CURRENT FILING DATE: 2005-01-12
; NUMBER OF SEQ ID NOS: 19247
; SEQ ID NO 112
; LENGTH: 1576
; TYPE: DNA
; ORGANISM: Escherichia coli K12
US-60-643-717-112

Alignment Scores:
Pred. No.: 2,108-313 Length: 1576
Score: 2039.00 Matches: 395
Percent Similarity: 99.7% Conservatives: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 82 Gaps: 0

US-10-673-786A-2 (1-396) x US-60-643-717-112 (1-1576)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 298 ATGATTGAGAACATTACCGCGCTCTCTGCCGACCGGATTTCTGGCGCTGGCGATCTGTT 357

Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 358 CGTCCGATGAACGTCGCGCAAAATTAACTCTCGGATTTGTTCTATAAGATGAGACG 417

Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 418 GGCNAACCCCGGTACTACCGAGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 477

Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 478 ACCACAAAATTAATCTCGCATTTGACGCGATCCCTGAAATTTGGTGGCTGCACTCAGAA 537

Qy 81 LeuLeuPheGlyLysGlySerAlaLeuLeuAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 538 CTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGTCGTCGACGCGACAGACT 597

Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 598 CGCGGGGCGACTGGCGCACTACGCGTGGCTGCGGATTTCTGGCAAAATACACGCGTT 657

Qy 121 LysArgValTropValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 658 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCATAGAGCGTCTTTAACTCTGCA 717

Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 718 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTTTGACTTCGAT 777

Qy 161 AlaLeuLysAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 983745

```
Db 778 GCACGTATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGTAGTGCTGTTCCATGGGTGC 837
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCCAACCGGTATCGACCCCTACGCTGGACCAATGGCAACACATGGCACAACCTC 897
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaIleValGlnGlyPheAlaArgGly 220
Db 898 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCGCTTACCAAGGGTTTGGCCCGGT 957
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 958 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGCGCTATGCAATAAAGAGACTGATTGTT 1017
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 1018 GCCAGTTCTACTCTAANAACCTTTGGCCTGTACAAACGAGCGGTGTTGGCGCTTGACTCTG 1077
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 1078 GTTGCTGCCGACAGTGAACCGGTTGATCGCGCAATTCAGCCAAATGAAGCGCGATTGCG 1137
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1138 GCTAACTACTCTAAACCCACACGACACACGCGCTTCTGTTGTGGCACCATCTCAGCAAC 1197
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1198 GATGCGTTACGTGGATTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT 1257
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 1258 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAGGCGCAAAACCGCGACTTCAGCTTT 1317
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1318 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCCCTGACAAAAGAACAAAGTCTCGGT 1377
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1378 CTGCGCGAAGAGTTGGCGTATATGCGGTTGCTTCTGCTCGGTAAATGTGGCGGAGT 1437
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1438 ACACCAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1485
```

RESULT 9

```
US-11-114-922-75
; Sequence 75, Application US/11114922
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 75
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
US-11-114-922-75
```

```
Alignment Scores: 3.87e-309 Length: 1191
Pred. No.: 2011.00 Matches: 390
Score:
```

```
Percent Similarity: 99.2% Conservative: 3
Best Local Similarity: 98.5% Mismatches: 3
Query Match: 98.3% Indels: 0
DB: 70 Gaps: 0
US-10-673-786A-2 (1-396) x US-11-114-922-75 (1-1191)
```

```
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTTACCGCGCTCTCTGCGGACCCGATTCTGGCGCTTGGCGATCTGTGTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 61 CGTGCCCATGAACGCTCCCGCAAAATTAACCTCGGATTTGGTCTCTATTACGATGAGACG 120
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 121 GGCAAAATCCCGGTACTGACCAGCGTGAAGGCTGAACAGATATCTGCTCGAAAAATGAA 180
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAACTTTACCTTCGGCATTTGACGGCATCCCTGAAATTTGGTCTGCTGCACTCAGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGACGGCACAGACT 300
Qy 101 ProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 301 CCGGGGGGCTCTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATAACACGCGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGCAACCATTAAGAGCGCTTTTAACTCTGCA 420
Qy 141 GlyLeuGluValArgGluTyrAlaTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 421 GGCTGGAAGTTGCGTAATACGCTTATATGATGCGGAAATACACTCTTTGACTTCGAT 480
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCACGTATTAACAGCCTGATGAAGCTCAGCGTGGCGACGTAGTGCTTCCATGGGTGC 540
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCCCAACCGGTATCGACCCCTACGCTGCGCAACATGGCAACACATGGCAACTC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTTCGCTTACCAGGGTTTGGCCCGTGT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 720
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTTCTACTCTAANAACCTTTGGCGCTGTACAAACGAGCGGTGTTGGCGCTTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGCTGCCGACAGTGAACCGTTGATGCGGCATTACGCCAAATGAAGCGCGCATTCGC 840
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 841 GCTAACTACTCTTCCCCCAGCACACGCGCTTCTGTTGTGGCCACCATCTCAGCAAC 900
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GATGCGTTACGTGGATTGGGAACCAAGAGCTGATGATGCGCCAGCGGTATTACAGCGT 960
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
```

Db 961 ATGGCTAGTTTCTGCTCAATACGCTGACGAGAAAGGCGCAACCGCGACTTTCAGCTTT 1020
Qy 341 ilelleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1021 ATCATCAACAGACAGCGCATGTTCTCTTCACTGCGCTGACAAAGACAACTGCTGGT 1080
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1081 CTGCGCAAGAGTTTGGCGTATATGCGTTGCTTCTGCTGCGTAAATGTGCGCGGATG 1140
Qy 381 ThrProAspAsnMetAlaProLeuGlyGluAlaIleValAlaValLeu 396
Db 1141 ACACAGATAAATGCTGCTGCGGAGCGATTTGTGCGAGTGTG 1188

RESULT 10

US-10-446-203-4902
; Sequence 4902, Application US/10446203
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/10/446,203
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: US/09/489,039
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4902
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-10-446-203-4902

Alignment Scores:

Pred. No.: 2,7e-296 Length: 1209
Score: 1931.00 Matches: 370
Percent Similarity: 96.5% Conservativity: 12
Best Local Similarity: 93.4% Mismatches: 14
Query Match: 94.4% Indels: 0
DB: 51 Gaps: 0

US-10-673-786a-2 (1-396) x US-10-446-203-4902 (1-1209)

Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 19 ATGTTTGAGAACATTACCGCGCCCGCCAGCCGCGATTTAGGTCTGCGCGATCTGTTT 78
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 79 CTGCGGATGACCGCCCTGAAATAATTAACCTCGGAATGGTGTACAGATGAACACC 138
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuGluAsnGlu 60
Db 139 GGTAAACCGCGTTCTGACAGCGTCAAAAAGCAGAGCAGTATCTGCTGAAAATGAA 198
Qy 61 ThrThrLysAsnTyrLeuGlyLysGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 199 ACGACTAAAAACTATCTGGCATCGATGATGTTCTCTGAAATTTGGTCTGCTGACCCAGGAG 258
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 259 CTGCTGTTTCGGTAAGGACATCGGATATTCGCTGATAACGCCCGCCAGCCGCGAGACA 318
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 319 CCCGCGGTACCGGTGCGTGTGCGCGAGACTTCTCTCGCCCAAAACACACCGACGTG 378
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 379 AAACGTGTGGTGAAGTAATCCGAGCTGCCGCAACCATAAAGCGGTATTACCTCTGCC 438

Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrPheAlaGluAsnHisThrLeuAspPheAsp 160
Db 439 GGGCTGGAAGTCGCGCAATACGCATACGACGCGGCTAACACCGCGCTGCACTTTGAT 498
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValLeuPheHisGlyCys 180
Db 499 GGTCTGCTGGCAGCGCTCAACGAAAGCCAGCGCGCGCGTGTACTGTTCACCGCTGC 558
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 559 TGCCACACCGCGCGGTATCGATCCGACGCTCGATCAGTGGCAGCAGCTGCGCGAGCTG 618
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 619 TCGGTGGAAGAGGCTGCTACCGCTGTTTCGCTTACCGAGGCTTCGCCCGCTG 678
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 679 CTGGAAGAGAGATGCTGAAGGCGTACCGCTTTTGTCTCTACATAAAGAGCTGCTGTGTC 738
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 739 GCCAGCTCTACTCGAAAAACTTTGGCTGTACACAGAGCGCTCGCGCGCTGCACCTG 798
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 799 GTCGCGCGGATCAGGAGACTGTAGACCGCGCTTCAGTCAGATGAAGTCGGTGTATCCGC 858
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValAlaAlaThrIleLeuSerAsn 300
Db 859 OCCAACTACTCGAACCCCGCTCGCGCATGCGCGCTCCGCTGCGCCACCATTCCTGAGCAAC 918
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 919 GATCGCTACCGCGCAATCTGGAGCAGGAATGACCGATATGCGCGCGCATCCAGCT 978
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGlyLysGlyAlaAsnArgAspPheSerPhe 340
Db 979 ATGGCTCTGCTGTTCTGCTCAATACCTGCGAGAGAAAGCGCGAGCGGACTTCAGCTTT 1038
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1039 ATCAGCCAGCAGAACCGCATGTTCTCATTCAGCGCGCTGACTAAAGACAGCGTGTGCGC 1098
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1099 CTGCGTGAAGAGTTCGCCCATCTATCGGTAGTCTCCGAGCGTATCAACGTGCGCGGATG 1158
Qy 381 ThrProAspAsnMetAlaProLeuGlyAlaIleValAlaValLeu 396
Db 1159 ACGCTGACATATGCGCGCGCTGTGCGAAGCCCATCGTCCGCTACTG 1206

RESULT 11

US-09-252-691-1606
; Sequence 1606, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691B
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 1606
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-09-252-691-1606

Alignment Scores:
Pred. No.: 3.91e-296 Length: 1209
Score: 1930.00 Matches: 370
Percent Similarity: 97.0% Conservativity: 14

Best Local Similarity: 93.4% Mismatches: 12
 Query Match: 94.4% Indels: 0
 DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691C-1606 (1-1209)

Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 19 ATGTTTGAGAACATTAATCTCGCGCTCTCGCGCAATTTCTGGCGCTGGCGATCTGTTT 78
 Qy 21 ArgAlaAaspGluAargProGlyLysIleAenLeuGlyLysValThrLysAaspGluThr 40
 Db 79 CGCGCCGACGACCGCCCTGGAAAAATCAACTGGGTATTTGGTGTATATAAGATGAAC 138
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
 Db 139 GGCAAACTCCGGTACTGACCGAGCTCAAAAAAGCTGAGCAGTATCTGCTGGAACGAA 198
 Qy 61 ThrThrLysAenTyrLeuGlyIleAaspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 199 ACCACCAAACTACCTCGGTATTTGATGATATCCCTGAAATTTGGTGTGCGTGCAC 258
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGlnThr 100
 Db 259 CTGCTGTTCCGCAAGGACGACCATTTGTAGCGGAAAAACGCGCCGACGCGCGAC 318
 Qy 101 ProGlyGlyThrGlyAlaLeuAargValAlaAaspPheLeuAlaLysAenThrSerVal 120
 Db 319 CCAGCGGTACCGGCACTCGCGGTGGCGACATTTCTGGCGAAAAACACATCTCTGTG 378
 Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 379 AAGCGTGTGGTAAAGCAATCCAAGCTGGCGCAACCATTAAGACGCTGTTAAATCTCGC 438
 Qy 141 GlyLeuGluValAargGluTyrAlaTyrTrpAlaGluAenHisThrLeuAaspPheAasp 160
 Db 439 GGTCTGGAAGTGCCTGTAATAGCTTACTACGACGACGCGCGCGCTGGATTTTCGAC 498
 Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGlyCys 180
 Db 499 GGCTGTGCGCAGCTGCTGCGGAGCGCGCGCGGCGATGTTGGTGTTCACGCGTGC 558
 Qy 181 CysHisAenProThrGlyIleAaspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
 Db 559 TGCCATAACCAACCGGTATCGATCTACGCTTACGAGTGGGAACAGCTGCCTAAGCTG 618
 Qy 201 SerValGluLysGlyTrpLeuProLeuPheAaspPheAlaTyrGlnGlyPheAlaAargGly 220
 Db 619 TCCGTTGAAAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
 Qy 221 LeuGluGluAaspAlaGluGlyLeuAargAlaPheAlaMetHisLysGluLeuIleVal 240
 Db 679 CTGGAAGNAGATGAGAAAGTCTGGCGCATTCGCGCGCTGATCAAGATTTGATGTC 738
 Qy 241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluAargValGlyAlaCysThrLeu 260
 Db 739 GCAAGCTCCTATTCGAAGAACTTTGGTCTGTATTAACGAGCGTGTGCGCGCTGTAC 798
 Qy 261 ValAlaAaspSerGluThrValAaspArgAlaPheSerGlnMetLysAlaAlaIleAarg 280
 Db 799 GTGGTGTGTCGAGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
 Qy 281 AlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
 Db 859 GCTAACTACTCCACCCGCGCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 918
 Qy 301 AspAlaLeuAargAlaIleTrpGluGlnGluLeuThrAaspMetArgGlnArgIleGlnArg 320
 Db 919 GATGCGTGTGCGGCATCTGGGAGCAAGAGCTGAACGATATGCGTCAAGCATTTTCG 978
 Qy 321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenAargAaspPheSerPhe 340
 Db 979 ATGCGTGTGCTGTTTGTGAACACGCTGCTGAGAAAGCGCTGACCGGTGACTTCAGCTTC 1038

Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuAarg 360
 Db 1039 ATCATCAAGCAGAACGCGATGTTCTCTTCAGCGGCTGACCAAGAGCAGGCTCTCGC 1098
 Qy 361 LeuAargGluGluPheGlyValTyrAlaValAlaSerGlyAargValAenValAlaGlyMet 380
 Db 1099 CTGCGTGAAGAGTTGGCGGTATGCGGTTCGATCTGCGCGCGTGAACGTTGCAAG 1158
 Qy 381 ThrProAaspAenMetalProLeuCysGluAlaIleValAlaValLeu 396
 Db 1159 ACGCCTGACAACTGCGCGCTGTGCGAAGCCATTGTCGCGGTGCTG 1206

RESULT 12
 US-09-252-691C-1606
 ; Sequence 1606, Application US/09252691C
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith G. Weinstein et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
 ; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.135
 ; CURRENT APPLICATION NUMBER: US/09/252.691C
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,145
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: US 60/074,787
 ; PRIOR FILING DATE: 1998-02-18
 ; NUMBER OF SEQ ID NOS: 11326
 ; SEQ ID NO 1606
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Enterobacter cloacae
 US-09-252-691C-1606

Alignment Scores:
 Pred. No.: 3,91e-296 Length: 1209
 Score: 1930.00 Matches: 370
 Percent Similarity: 97.0% Conservative: 14
 Best Local Similarity: 93.4% Mismatches: 12
 Query Match: 94.4% Indels: 0
 DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691C-1606 (1-1209)

Qy 1 MetPheGluAenIleThrAlaAProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
 Db 19 ATGTTTGAGAACATTAATCTCGCGCTCTCGCGCAATTTCTGGCGCTGGCGATCTGTTT 78
 Qy 21 ArgAlaAaspGluAargProGlyLysIleAenLeuGlyLysValThrLysAaspGluThr 40
 Db 79 CGCGCCGACGACCGCCCTGGAAAAATCAACTGGGTATTTGGTGTATATAAGATGAAC 138
 Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAenGlu 60
 Db 139 GGCAAACTCCGGTACTGACCGAGCTCAAAAAAGCTGAGCAGTATCTGCTGGAACGAA 198
 Qy 61 ThrThrLysAenTyrLeuGlyIleAaspGlyIleProGluPheGlyArgCysThrGlnGlu 80
 Db 199 ACCACCAAACTACCTCGGTATTTGATGATATCCCTGAAATTTGGTGTGCGTGCAC 258
 Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGlnThr 100
 Db 259 CTGCTGTTCCGCAAGGACGACCATTTGTAGCGGAAAAACGCGCCGACGCGCGAC 318
 Qy 101 ProGlyGlyThrGlyAlaLeuAargValAlaAaspPheLeuAlaLysAenThrSerVal 120
 Db 319 CCAGCGGTACCGGCACTGCGCGTGGCGCAGATTTCTGGCGAAAAACACATCTCTGTG 378
 Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
 Db 379 AAGCGTGTGGTAAAGCAATCCAAGCTGGCGCAACCATTAAGACGCTGTTAAATCTCGC 438
 Qy 141 GlyLeuGluValAargGluTyrAlaTyrTrpAlaGluAenHisThrLeuAaspPheAasp 160

```
Db 439 GGTCTGGAAGTGGTGAATACCCCTACTACGACGCGCCAGCCGCGTGGATTTCGAC 498
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 499 GGCCTGCTGGCCAGCTGAGCGAAGCCAGGCGGGGATGTGTCTGTTTCCACGGCTGC 558
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 559 TGCCATAACCCCAACCGGTATCGATCCTACGCTTGAAGAGTGGGAACAGCTGGCTAAGCTG 618
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 619 TCCGTTGAAAAGGTGGCTGGCGCTGTGTGATCTTCGCTTACAGGGCTTGGCCGCTGGT 678
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 679 CTGGAAGAGATGCAAGAGTCTGGCGCATTCGGCGCGCTGCATCAAGAGTTGTATGTC 738
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValaCysThrLeu 260
Db 739 GCAAGCTCTCTATTCGAAGAACTTTGGTCTGTATTAACGAGCGTGTGGCGCTGTACGCTG 798
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 799 GTGGCTGCTGACGAAGCAACCGTGCATCGCGCATTCAGCCAGATGAAGTCCGTGATCGCG 858
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 859 GCTAACTACTCCAAACCCAGCCGACACGCGTGTCTGTGGTGGCGACCATCTGAGCAAC 918
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
Db 919 GATGGCTGGCGGCCATCTGGAGGACAGAGCTGAACGATATGCGTCAAGCGCATTCAGCGC 978
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 979 ATGCGTGTCTGCTTTGTGAACACGCTGGCTGAGAAAGCGCTGACCGTACCTCAGCTTC 1038
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluValLeuArg 360
Db 1039 ATCATCAGCAGAAACGGCATGTCTCTTCAGCGCGCTGACCAAGAGAGAGGTGCTGCGC 1098
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1099 CTGCGTGAAGAGTTTGGCGTGTATGCCGTTGATCTGGCCGCGTGAACGTTGACGTATG 1158
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1159 AGCCCTGACAAATGGCGCGCTGTGCGAAGCAATGTGCGCGTGTCTG 1206
```

```
RESULT 13
US-10-417-886-1606
; Sequence 1606, Application US/10417886
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/10/417,886
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US/09/252,691C
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/094,145
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 60/074,787
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 11326
; SEQ ID NO 1606
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-10-417-886-1606
```

```
Alignment Scores: 3.91e-296 Length: 1209
Pred. No.: 1930.00 Matches: 370
Score: 97.0% Conservative: 14
Percent Similarity: 93.4% Mismatches: 12
Best Local Similarity: 94.4% Indels: 0
Query Match: 51 Gaps: 0
US-10-673-786A-2 (1-396) x US-10-417-886-1606 (1-1209)
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 19 ATGTTTGAACAATTTACTGCGCGCTCTCCGCCACCAATTTCTGGGCTGGCCGATCTGTTT 78
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 79 CGCGCCGACGACCGCCCTGGAAAATCAACTGGGTATTTGGTGTATATTAAGATGAACACC 138
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuGluAsnGlu 60
Db 139 GGCAAAACCTCCGCTACTGACCAGCGTCAAAAAGCTGAGCAGTATCTGCTGGAACACGAA 198
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 199 ACCACCAAAAACTACTCGGTATTTGATGTTATCCCTGAATTTGGTCTGCACCCAGGAG 258
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 259 CTGCTGTTTCGGCAAGGAGGAGCACCATTGTGAGCGAAAACCGCCGCCGCGGACGACC 318
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 319 CCAGCGGTACCGGGCGACTCGCGCTGGCGGAGATTTCTCTGGCGAAACACATCTCTGTG 378
Qy 121 LysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 379 AAGCGTGTGGTAAAGCAATCCAAAGCTGGCGCAACCAATAAGAGCGTGTAAATTTCTGCG 438
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 439 GGTCTGGAAGTGGCGTGAATACGCTTACGACGCGCAGCCAGCCGCGCTGTGATTTTCGAC 498
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 499 GGCCTGCTGGCCAGCTGAGCGAAGCCAGCGGGGAGTGTGTGCTGTCTCCACGGCTGC 558
Qy 181 CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 559 TGCCATAACCCCAACCGGTATCGATCCTACGCTTGAAGAGTGGGAACAGCTGGCTAAGCTG 618
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 619 TCCGTTGAAAAGGTGGCTGGCGCTGTGTGATCTTCGCTTACAGGGCTTGGCCGCTGGT 678
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 679 CTGGAAGAGATGCAAGAGTCTGGCGCATTCGGCGCGCTGCATCAAGAGTTGTATGTC 738
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyValaCysThrLeu 260
Db 739 GCAAGCTCTCTATTCGAAGAACTTTGGTCTGTATTAACGAGCGTGTGGCGCTGTACGCTG 798
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 799 GTGGCTGCTGACGAAGCAACCGTGCATCGCGCATTCAGCCAGATGAAGTCCGTGATCGCG 858
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 859 GCTAACTACTCCAAACCCAGCCGACACGCGTGTCTGTGGTGGCGACCATCTGAGCAAC 918
Qy 301 AspAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
```

```
Db 919 GATGGCTGCGCGCATCTGGGAGCAAGAGCTGAACGATATGCGTCAGCGCAATTCAGCGC 978
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 979 ATGGCTGCTGTTGTGTGAACACGCTGGCTGAGAAAGCGCTGACCGTGACTTCAGCTTC 1038
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1039 ATCATCAAGCAGAACGGCATGTTCTCTTCAGCGGCTGACCAAGAGCAGGTGCTGCGC 1098
Qy 361 LeuArgGlnGluPheGlyValThrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1099 CTGGCTGAAGAGTTGGCGTGATGCGCGTTCATCTGGCGCGTGAACGTTGCGAGGTATG 1158
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1159 ACGCTGACAACATGGCGCGCTGTGCGAGCCATTTGCGCGGTGCTG 1206

RESULT 14
US-09-252-691C-1616/c
; Sequence 1616, Application US/09252691C
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; FILE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691B
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 1616
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-09-252-691-1616

Alignment Scores:
Pred. No.: 3,97e-296 Length: 1218
Score: 1930.00 Matches: 370
Percent Similarity: 97.0% Conservative: 14
Best Local Similarity: 93.4% Mismatches: 12
Query Match: 94.4% Indels: 0
DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691-1616 (1-1218)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1200 ATGTTTGAGAACATTAATCTGCGCGCTCCTGCGGACCAATTTCTGGCGCTGGCGCATCTGTTT 1141
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValThrLysAspGluThr 40
Db 1140 CGCGCCGAGACCGCCCTGGAAAATCACTGGGTATGGTGATATTAAGATGAAC 1081
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnThrLeuLeuGluAenGlu 60
Db 1080 GGCAAACTCCGGTACTGACCAAGCTGCAAAAGCTGAGCAGTATCTGCTGGAACGAA 1021
Qy 61 ThrThrLysAenThrLeuGlyLysAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 1020 ACCACCAAACTACTCGGTATGATGGTATCCCTGAAATTTGGTCGTCGACCCAGAG 961
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 960 CTGCTGTTCCGCAAGGAGGAGCACCATTGTGAGCGAAAACCGCGCGGACGCGCGCAGACC 901
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 900 CCAGCGGTACCGCGCATCTGCGGTATGCGCGCAGATTTCTTGGCGAAAACACATCTCTGTG 841
Qy 121 LysArgValThrValSerAsnProSerThrProAsnHisLysSerValPheAsnSerAla 140
Db 840 AAGCGTGTGTGGTAAAGCAATCCAAGCTGGCGGACCAACCAAGAGCGGTGTTTAATCTGCG 781
```

```
Qy 141 GlyLeuGluValArgGluThrAlaThrTyrAspAlaGluAenHisThrLeuAspPheAsp 160
Db 780 GGTCTGGAAGTGGGTGAATACGCTACTACGACGCAAGCCACGCGCTGGATTTGCGAC 721
Qy 161 AlaLeuIleAsnSerLysAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 720 GGCTGTGTCGCCAGCTGAGCGAAGCCAGCGGCGCATGTGTGCTGTTCACCGGTGTC 661
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnThrLeuAlaGlnLeu 200
Db 660 TGCCATCAACCAACCGGTATCGATCTACGCTTGAGCAGTGGGAACAGCTGGCTAAGCTG 601
Qy 201 SerValGluLysGlyThrLeuProLeuPheAspPheAlaThrGlnGlyPheAlaArgGly 220
Db 600 TCGCTTGAAAAGAGCTGGCTGCGCTGTTTGTACTTCGCTTACCAGGCGCTTTGCGCGGT 541
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 540 CTGGAAGAAGATGAGAAGGTCTGCGCGCAATTCGCCCGCTGCATCAAGAGTTGATAGTC 481
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 480 GCAAGCTCTCTATTTCCAAGAACTTTGGTCTGTATAACGAGCGTGTGCGCGCTGTACGCTG 421
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 420 GTGGCTGCTGCAAGCAACACCGTCGATCGCGCATTCAGCCAGATGAAGTCGCTGATCCGC 361
Qy 281 AlaAsnTyrSerAsnProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 360 GCTAACTACTTCCAAACCCACCGCACAGGTGCGTCTGTGTCGCGAGCATCTGAGCAAC 301
Qy 301 AspAlaLeuArgAlaIleThrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 300 GATGCGCTGCGCGCATCTGGGAGCAAGAGCTGAACGATATGCGTCAGCGCATTCAGCGC 241
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 240 ATGCGCTGCTGTTGTGAACACGCTGGCTGCGTGAAGAGCGCTGACCGCTGACTTCAGCTTC 181
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 180 ATCATCAAGCAGAACCGCATGTTCTCTTCCAGCGGCTGACCAAGAGCAGGTGCTGCGC 121
Qy 361 LeuArgGluGluPheGlyValThrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 120 CTGGTGAAGAGTTTGGCGTGATGCGGTGATGCGTGTGATCTGCGCGCGTGAACGTTGCAAGGTATG 61
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 60 ACGCTGACAACATGGCGCGCTGTGCGAAGCCATTTGCGCGGTGCTG 13
```

```
RESULT 15
US-09-252-691C-1616/c
; Sequence 1616, Application US/09252691C
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; FILE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691C
; PRIOR FILING DATE: 1999-02-18
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-07-24
; PRIOR FILING DATE: 1998-07-24
; PRIOR FILING DATE: 1998-07-24
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 11326
; SEQ ID NO 1616
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-09-252-691C-1616
```

Alignment Scores:

Pred. No.: 3.97e-296 Length: 1218
Score: 1930.00 Matches: 370
Percent Similarity: 97.0% Conservative: 14
Best Local Similarity: 93.4% Mismatches: 12
Query Match: 94.4% Indels: 0
DB: 22 Gaps: 0

US-10-673-786A-2 (1-396) x US-09-252-691C-1616 (1-1218)

```
Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1200 ATGTTTGAGAACATTACTCGCGCTCTCTGCCACCAATCTGGGCTGGCGCATCTGTTT 1141
Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrLysAspGluThr 40
Db 1140 CGCGCCGACGACCGCCCTGGAAAATCAACCTGGGTATGGTGATATATAAGATGAAC 1081
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGlu 60
Db 1080 GGC AAAA CTCCGGTACTACCGAGCGTCAAAAAGCTGAGCAGTATCTCTGGA AAAA CGAA 1021
Qy 61 ThrThrLysAsnTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 1020 ACCACCAAAAATCTACCTCGGTATTGATGGTATCTCCGTAATTTGGTCGTGCACCCAGGAG 961
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThr 100
Db 960 CTGCTGTTCCGCAAGGAGGAGCACCATTGTGAGCGAAAACCGGCCCGCCAGCGGAGACC 901
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 900 CCAGCGGTACCGCGGCACCTCGCGCTGCGGCAGATTTCTTGGCGAAAACACATCTCTGTG 841
Qy 121 LysArgValTyrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 840 AAGCGTGTGGGTAAAGCAATCAACAGCTGGCGGAAACCAATAAGAGCGGTGTTTAATTCGCG 781
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 780 GGTCTGGAAGTGGTGAATAGCCCTACTACGACGCGCAGCCAGCCGCTGGATTTTCGAC 721
Qy 161 AlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 720 GGCTGTCTGGCAGCGCTGAGCGAAGCCAGCGGCGGATGTGGTCTGTTCACCGGCTGC 661
Qy 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 660 TGCCATAACCCAAACCGGTATCGATCCTACGCTTGAGCAGTGGGAACAGCTGGCTAAGCTG 601
Qy 201 SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 600 TCCGTTGAAAAGGTGGCTGCGCTGTTGACTTCGCTTACCGAGGCTTTGGCCCGGTGT 541
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysLysLeuIleVal 240
Db 540 CTGGAGAGAGATGAGAGAGGTCTGGCGCATTCGCCCGCGTGATCAAGAGTTGATGTC 481
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 480 GCAAGCTCTATTCCAAAGACTTTGGTCTGTATACGAGCGGTGTCGGCGCTGTACGCTG 421
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 420 GTGGCTGTCTGCAAGACCGTTCGATCGCATTCGCGATTCAGCCAGATGAAGTCGTGATCCGC 361
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 360 GCTAACTACTCCAAACCCAGCGCACCGGTGGTCTGTGGTCGCGACCATCTCTGAGCAAC 301
Qy 301 AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 300 GATCGCTGCGGCCATCTGGGAGCAAGAGCTGAACGATATGCGTTCAGCGCATTCAGCGC 241
```

Search completed: March 15, 2006, 17:48:21

Job time : 8035 secs

```
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe 340
Db 240 ATGCGTCTGCTGTTTGTGAACACACGCTGGCTGAGAAAGCGCTGACCGTGACTTCAGCTTC 181
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 180 ATCATCAAGCAGAACCGCATGTTCTCTTCAGCGGCTTGACCAAGAGCAGGTGCTGCGC 121
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 120 CTGCGTGAAGAGTTTGGCGTGTATGCCGCTTGCACTCGCCGCTGAACGTTGACGATG 61
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 60 ACGCTGACAACTGGCGCGCTGTCGGAAGCCATTGTGCGCGTGTCTG 13
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2006, 15:45:30 ; Search time 410 Seconds
(without alignments)
2682.737 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045
Sequence: 1 MFENITAPADPILGLADLF.....VAGMTPDNMPLCEAIVAVL 396

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6590814 seqs, 1388790629 residues

Total number of hits satisfying chosen parameters: 13181628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10673786/runat_14032006_135618_22781/app_query.fasta_1
-DB=pending_Patents_NA_New -QFMT=fastap -SUFFIX=trnp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss03h -USER=US10673786 @CGN 1 1 70 @runat_14032006_135618_22781
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents_NA_New.*

```
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US12_NEW_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US13_NEW_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US14_NEW_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US15_NEW_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US16_NEW_COMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2045	100.0	1191	10	US-11-302-262-19
2	2045	100.0	1241	1	PCT-US05-47111-4
3	2045	100.0	1241	12	US-60-752-355-4
4	2039	99.7	1576	1	PCT-US06-00964-112
5	2039	99.7	1576	8	US-11-330-403-112
6	1310	64.1	2242716	7	US-10-018-4708-1

C

7	913	44.6	1230	1	PCT-US05-47111-133	Sequence 133, App
8	913	44.6	1230	12	US-60-752-355-133	Sequence 133, App
9	909	44.4	1224	1	PCT-US05-47111-446	Sequence 446, App
10	909	44.4	1224	12	US-60-752-355-446	Sequence 446, App
11	885	43.3	1668	7	US-10-953-349-16412	Sequence 16412, A
12	865	42.3	1885	12	US-60-771-906-33	Sequence 33, Appl
13	814	39.8	1218	10	US-11-299-286-563	Sequence 563, App
14	806	39.4	1301	7	US-10-953-349-11025	Sequence 11025, A
15	794	38.8	2339	8	US-11-329-121-902	Sequence 902, App
16	794	38.8	2339	12	US-60-741-050-148	Sequence 148, App
17	794	38.8	2339	12	US-60-763-366-696	Sequence 696, App
18	794	38.8	2417	8	US-11-329-121-903	Sequence 903, App
19	794	38.8	2417	12	US-60-763-366-697	Sequence 697, App
20	794	38.8	2465	8	US-11-329-121-901	Sequence 901, App
21	794	38.8	2465	12	US-60-763-366-695	Sequence 695, App
22	792	38.7	2411	10	US-11-266-748A-25190	Sequence 25190, A
23	783	38.3	2005	10	US-11-266-748A-184254	Sequence 184254, A
24	783	38.3	2005	10	US-11-266-748A-191752	Sequence 191752, A
25	759	37.1	2365	10	US-11-266-748A-68736	Sequence 68736, A
26	759	37.1	2365	10	US-11-266-748A-121547	Sequence 121547, A
27	758	37.1	1293	10	US-11-299-286-769	Sequence 769, App
28	758	37.1	1612	7	US-10-953-349-3641	Sequence 3641, Ap
29	736	36.0	1937	10	US-11-224-525-639	Sequence 639, App
30	736	36.0	1937	10	US-11-224-525-639	Sequence 639, App
31	735.5	36.0	1299	1	PCT-US05-47111-2	Sequence 2, Appli
32	735.5	36.0	1299	12	US-60-752-355-2	Sequence 2, Appli
33	726.5	35.5	1230025	8	US-11-366-942-1	Sequence 1, Appli
34	652.5	31.9	1373	10	US-11-266-748A-68993	Sequence 68993, A
35	652.5	31.9	1373	10	US-11-266-748A-105184	Sequence 105184, A
36	652.5	31.9	1373	10	US-11-266-748A-121804	Sequence 121804, A
37	606	29.6	2777	10	US-11-266-748A-68992	Sequence 68992, A
38	606	29.6	2777	10	US-11-266-748A-105183	Sequence 105183, A
39	606	29.6	2777	10	US-11-266-748A-121803	Sequence 121803, A
40	598.5	29.3	1356	1	PCT-US05-47111-3	Sequence 3, Appli
41	598.5	29.3	1356	12	US-60-752-355-3	Sequence 3, Appli
42	594.5	29.1	1045	10	US-11-266-748A-9142	Sequence 9142, Ap
43	594.5	29.1	1045	10	US-11-266-748A-64005	Sequence 64005, A
44	594.5	29.1	1045	10	US-11-266-748A-66837	Sequence 66837, A
45	494	24.2	801	8	US-11-360-355-48767	Sequence 48767, A

ALIGNMENTS

RESULT 1

```
US-11-302-262-19
; Sequence 19, Application US/11302262
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: A method of improving growth and yield of plants under reduced
; FILE REFERENCE: OP05262
; CURRENT APPLICATION NUMBER: US/11/302,262
; CURRENT FILING DATE: 2005-12-14
; PRIOR APPLICATION NUMBER: JP 2003-198559
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1188)
US-11-302-262-19
```

Alignment Scores:

Pred. No.:	2e-237	Length:	1191
Score:	2045.00	Matches:	396
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-673-786A-2 (1-396) x US-11-302-262-19 (1-1191)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 1 ATGTTTGAGAACATTACCGCGCTCTCTGCCGACCGGATTCTGGCGCTGGCGATCTGTTT 60
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValThrLysAspGluThr 40
Db 61 CGTCCGATGAACGTCCTGGCAAAATTAACTCGGATTGGTGCTATAAAGATGAGAGC 120
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnThrLeuLeuGluAenGlu 60
Db 121 GGCAAAACCCCGGATCTACCGAGCGTGAAAGGCTGAACGATCTGCTCGAAATGAA 180
Qy 61 ThrThrLysAenThrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 181 ACCACCAAAATTACCTCGGATTGACGGCATCTCTGAAATTGGTGGCTGCACCTCAGGAA 240
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 241 CTGCTGTTTGGTAAAGGTAGCGCTGATCAATGACAAACGCTGCTCGCAGCGCACAGCT 300
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 301 CCGGGGGGACCTGGCGCATCTACGGTGGCTGCCGATTCTCGGCAAAATAACACGCGTT 360
Qy 121 LysArgValTrpValSerAsnProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 361 AAGCGTGTTGGTGAGCAACCCAGCTGGCGGACCCATAGAGCGTCTTTAACTCTGCA 420
Qy 141 GlyLeuGluValArgGluThrAlaThrLysAspAlaGluAenHisThrLeuAspPheAsp 160
Db 421 GGTCTGGAAGTTCTGTAATACGCTATTATGATGCGGAAATACACACTCTTTGACTTCGAT 480
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 481 GCACGTATTAAACAGCCTGAATGAAGCTCAGGCTGGCGAGTAGTGCTGTTCCATGGCTGC 540
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu 200
Db 541 TGCCATAACCAACCGGTATCGACCTACGCTGGACCAATGGCAACACTGGCAACTC 600
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaThrGlnGlyPheAlaArgGly 220
Db 601 TCCGTGTGAGAAGGCTGCTTACCGCTGTTGACTTCGCTTACCGAGGTTTTTCCCGGCT 660
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleVal 240
Db 661 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTTCGGGCTATGCTATAAAGAGCTGATTGTT 720
Qy 241 AlaSerSerThrSerLysAenPheGlyLeuThrAenGluArgValGlyAlaCysThrLeu 260
Db 721 GCCAGTTCCTACTCTAAAACTTTGGCCGTGTACCAACGAGCGTGTGGCGCTTGACTCTG 780
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
Db 781 GTTGCTCGGCACAGTGAACCGCTTGATCGCGCATTCAGCCAAATGAAAGCGCGATTTCG 840
Qy 281 AlaAenThrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
Db 841 GCTAACTACTTAACCCACAGCACACGGCGCTTCTGTTGTTGGCACCATCTCTGAGCAAC 900
Qy 301 AspalAlaLeuArgAlaIleTrpGluGlnLeuThrAspMetArgGlnArgIleGlnArg 320
Db 901 GATCGTTTACGTGGATTGGGAACAAGAGCTGATGATGGCCAGCGTATTTCAGCGT 960
Qy 321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenArgAspPheSerPhe 340
Db 961 ATGCGTCAGTTGTTTCGTAATACGCTGCAGGAAAGGCGCAACCCGCACTTCAGCTTT 1020
Qy 341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360

Db 1021 ATCATCAACAGACGCGATGTTCTCTCAGTGGCTGACAAAAGAACAGTCTCGGT 1080
Qy 361 LeuArgGluGluPheGlyValThrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
Db 1081 CTGCGCGAAGAGTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTGCGCGGATG 1140
Qy 381 ThrProAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1141 ACACCAAGATAACATGGCTCGCTGTGCGAAGCGATTGTGGCAGTGCTG 1188

RESULT 2

PCT-US05-47111-4
; Sequence 4, Application PC/TUS0547111
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: PCT/US05/47111
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Escherichia coli
PCT-US05-47111-4

Alignment Scores:
Pred. No.: 2,12e-237 Length: 1241
Score: 2045.00 Matches: 396
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x PCT-US05-47111-4 (1-1241)

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 25 ATGTTTGAGAACATTACCGCGCTCTCTGCCGACCGGATTCTGGCGCTGGCGATCTGTTT 84
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValThrLysAspGluThr 40
Db 85 CGTCCGATGAACGTCCTGGCAAAATTAACTCGGAGTTGGTGCTATAAAGATGAGAGC 144
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluGlnThrLeuLeuGluAenGlu 60
Db 145 GGCAAAACCCCGGATCTGACCGCGTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 204
Qy 61 ThrThrLysAenThrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 205 ACCACCAAAATTTACCTCGGCAITTCGCGCATCCCTGAATTTGGTCTGCTCAGTCA 264
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 265 CTGCTGTTTGGTAAAGGTAGCGCTGATCAATGACAAACGCTGCTCGCAGCGCACACT 324
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 325 CCGGGGGGCACTGGCGCATCTACCGTGGCTGCCGATTCTCGCAAAATAACACGCGTT 384
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 385 AAGCGTGTGGGTGAGCAACCCAGCTGGCCGACCATTAAGAGCGCTTTTAACTCTGCA 444
Qy 141 GlyLeuGluValArgGluThrAlaThrLysAspAlaGlnAenHisThrLeuAspPheAsp 160
Db 445 GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAATACACTCTTTGACTTCGAT 504
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 505 GCACTGATTAAACGCTGAATGAAGCTCAGGCTGGCGACGCTAGTGCTGTTCCATGGCTGC 564

181	Qy	CysHisAsnProThrGlyIleAspProThrLeuGluGlnTrpGlnThrLeuAlaGlnLeu	200
565	Db	TGCCATAACCCAAACCGGTATCGACCTACGCTGGAAACAATGGCAAAACACTGGCAACAATC	624
201	Qy	SerValGluLysGlyTyrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly	220
625	Db	TCCGTTGAGAAAGCGCTGTTACCGCTGTTGACTTCGCTTACGAGGGTTTGGCCGTGGT	684
221	Qy	LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal	240
685	Db	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT	744
241	Qy	AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu	260
745	Db	GCCAGTTCCTACTCTAAAAACCTTTGGCTGTGACAAACGAGCGTGTGGCGCTTGTACTCTG	804
261	Qy	ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg	280
805	Db	GTTGCTCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTTCG	864
281	Qy	AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn	300
865	Db	GCTAACTACTCTAAACCCACGACACACGGCGCTTCTGTTGTCACCATCTCTGAGCAAC	924
301	Qy	AspAlaLeuArgAlaIleTyrGluGlnGluLeuThrAspMetArgGlnArgGleGlnArg	320
925	Db	GATCGGTTACGTGCATTTGGAAACAAGAGCTGACTGATATGCCCCAGCGTATTACGCGT	984
321	Qy	MetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAspPheSerPhe	340
985	Db	ATGCGTCAGTTGTTTCGTCAATACGCTGCAGGAAAGCGCAAAACCGGACTTTCAGCTTT	1044
341	Qy	IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg	360
1045	Db	ATCATCAAAACAGAACCGCATGTTCTCCCTTCAGTGGCTGTGCAAAAGAACCAAGTGCCTCGT	1104
361	Qy	LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet	380
1105	Db	CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTCGCTAAATGTGGCGGGATG	1164
381	Qy	ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu	396
1165	Db	ACACCATCAACATGGCTCGGCTGTGGAAACGCAATGTGCAGTGCCTG	1212

```

RESULT 3
US-60-752-355-4
; Sequence 4, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Escherichia coli
US-60-752-355-4

Alignment Scores:
Pred. No.:                2.12e-237
Score:                    2045.00
Percent Similarity:      100.0%
Best Local Similarity:   100.0%
Query Match:             100.0%
DB:                       12

Length: 1241
Matches: 396
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-673-786A-2 (1-396) x US-60-752-355-4 (1-1241)

Qy      1  MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20

```

```
Qy 381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1165 ACACAGATAACATGGCTCCGCTGTCGCAAGCGATTGTGGCAGTGCTG 1212

RESULT 4
PCT-US06-00964-112
; Sequence 112, Application PC/TUS0600964
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/00964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 112
; LENGTH: 1576
; TYPE: DNA
; ORGANISM: Escherichia coli K12
PCT-US06-00964-112

Alignment Scores:
Pred. No.: 1,63e-236 Length: 1576
Score: 2039.00 Matches: 395
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 1 Gaps: 0

US-10-673-786A-2 (1-396) x PCT-US06-00964-112 (1-1576)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 298 ATGATTGAGAACATTACCGCGCTCTCCGACCGGATTCTGGCGCTGGCGGCTGTTT 357
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 358 CGTCCGATGAACGTCGCGGCAAAATTAACTCCGGATTGGTGTCTATAAAGATGAGACG 417
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 418 GCGAAACCCCGGTACTGACGAGCGTGAAAGGCTGAACAGTATCTCTCGAAATGAA 477
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 478 ACCACAAAATTAACCTCGGCATTGACGCGCATCCCTGAATTTGGTCTGCACTCAGGAA 537
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 538 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGCAGACT 597
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 598 CCGGGGGGACCTGGCGCACTAGCGGTGCTGCCGATTCTCGGCAAAAATAACACGCTT 657
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 658 AAGCGTGTGGTGAGCAACCCAGCTGCGCGCAACCAATAGAGCGTCTTTAACTCTGCA 717
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAspPheAsp 160
Db 718 GGTCTGGAAGTTCTGTGAATAGCTTATATGATGCGGAAATTCACACTCTTGACTTCGAT 777
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 778 GCACGTGATTAAACGCTGAAAGAGCTCAGGCTGGCGACGATGCTGTTTCATGGCTGC 837
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCCAACCGGTATCGACCTTAGCTGGGAACAATGGCAACACTGGCACAAC 897
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 941 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60

898 TCCGTTGAGAAAGCGTGGTTACCGCTGTTTACCTTCGCTTACCAGGGTTTTTGGCCGCTGGT 957
221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaMetHisLysGluLeuIleVal 240
958 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCGCGCTATGCATAAAGAGCTGATTGTT 1017
241 AlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCysThrLeu 260
1018 GCCAGTTTCTTACTCTAAAAACTTTGGCCTGTACAAACGAGCGGTGTGGCGCTTGACTCTG 1077
261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArg 280
1078 GTTGCTCCGACACAGTGAACCGTTGATCGCGCATTCAGCCAAATGMAAGCGCGATTGCGC 1137
281 AlaAenTyrSerAenProProAlaHisGlyAlaSerValValAlaThrIleLeuSerAen 300
1138 GCTAACTACTCTAACCCACCAGCACACGCGCTCTCTGTGTGTGGCCACCATCTCAGCAAC 1197
301 AspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgGlnArg 320
1198 GATCGCTTACGTGCGATTGGGAACAAGAGCTGACTGATATGCCGCGAGCGTATTTCAGCGT 1257
321 MetArgGlnLeuPheValAenThrLeuGlnLysGlyAlaAenArgAspPheSerPhe 340
1258 ATGCGTCAGTTGTTCTGTCATATACGCTCAGGAAAGGCGCAACCGCGACTTCAGCTTT 1317
341 IleIleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
1318 ATCATCAACAGAACGCGCATGTTCTCTTCAGTGGCTGACAAAGAAACAAGTGTGCGT 1377
361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMet 380
1378 CTGCGCGAAGAGTTGGCGTATATGCGGTGCTTCTCGTAAATGTGGCGGGAATG 1437
381 ThrProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
1438 ACACAGATAACATGGCTCCGCTGTCGCAAGCGATTGTGGCAGTGCTG 1485

RESULT 5
US-11-330-403-112
; Sequence 112, Application US/11330403
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 112
; LENGTH: 1576
; TYPE: DNA
; ORGANISM: Escherichia coli K12
US-11-330-403-112

Alignment Scores:
Pred. No.: 1,63e-236 Length: 1576
Score: 2039.00 Matches: 395
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 8 Gaps: 0

US-10-673-786A-2 (1-396) x US-11-330-403-112 (1-1576)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 298 ATGATTGAGAACATTACCGCGCTCTCCGACCGGATTCTGGCGCTGGCGGCTGTTT 357
Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 358 CGTCCGATGAACGTCGCGGCAAAATTAACTCCGGATTGGTGTCTATAAAGATGAGACG 417
Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 418 GCGAAACCCCGGTACTGACGAGCGTGAAAGGCTGAACAGTATCTCTCGAAATGAA 477
Qy 61 ThrThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGlnGlu 80
Db 478 ACCACAAAATTAACCTCGGCATTGACGCGCATCCCTGAATTTGGTCTGCACTCAGGAA 537
Qy 81 LeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGlnThr 100
Db 538 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGCGCAGACT 597
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSerVal 120
Db 598 CCGGGGGGACCTGGCGCACTAGCGGTGCTGCCGATTCTCGGCAAAAATAACACGCTT 657
Qy 121 LysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSerAla 140
Db 658 AAGCGTGTGGTGAGCAACCCAGCTGCGCGCAACCAATAGAGCGTCTTTAACTCTGCA 717
Qy 141 GlyLeuGluValArgGluTyrAlaTyrTrpAspAlaGluAenHisThrLeuAspPheAsp 160
Db 718 GGTCTGGAAGTTCTGTGAATAGCTTATATGATGCGGAAATTCACACTCTTGACTTCGAT 777
Qy 161 AlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 778 GCACGTGATTAAACGCTGAAAGAGCTCAGGCTGGCGACGATGCTGTTTCATGGCTGC 837
Qy 181 CysHisAenProThrGlyIleAspProThrLeuGluGlnTyrGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCCAACCGGTATCGACCTTAGCTGGGAACAATGGCAACACTGGCACAAC 897
Qy 201 SerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 941 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
```

```

Db 418 GCGCAAAACCCGGTACTGACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 477
Qy 61 ThrThrLysAsnTyrLeuGlyLeuAspGlyLeuProGluPheGlyArgCysThrGlnGlu 80
Db 478 ACCACAAAATTTACTCTGGCATTGACGGCATCCCTGAAATTTGGCTGCGTACTCAGAA 537
Qy 81 LeuLeuPheGlyLysSerAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 538 CTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGCTGCTCGACGCGCACAGACT 597
Qy 101 ProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerVal 120
Db 598 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGATTTCTTGGCAAAAATACCAACGCTT 657
Qy 121 LysArgValTrrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAla 140
Db 658 AAGCGTGTGTGGTGAGCAACCAAGCTGGCCGAACCAATGAAGGCTCTTTAACTCTGCA 717
Qy 141 GlyLeuGluValArgGlyLysAlaTyrTyrAspAlaGluAsnHisThrLeuAspPheAsp 160
Db 718 GGTCTGGAAGTTCTGTAATACCTTATTATGATGCGGAATACACACTCTTCACTTCGAT 777
Qy 161 AlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGlyCys 180
Db 778 GCACCTGATTAAACAGCTCAATGAAGCTCAGGCTGCGAGTAGTGTGTTCATGCTGCTGC 837
Qy 181 CysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrrGlnThrLeuAlaGlnLeu 200
Db 838 TGCCATAACCCCAACCGGTATCACCTACGCTGGAACAATGCAACAACACTGGCAACATC 897
Qy 201 SerValGluLysGlyTrrLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGly 220
Db 898 TCCGTTGAGAAAGCTGTGTTACCGCTGTTTGTACTTACGAGGTTTTCGCCGTGTGT 957
Qy 221 LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuVal 240
Db 958 CTGGAAGAAGATGCTGAAGCACTGCGCGCTTTTCGCGCTATGCATAAAGAGCTGATTGT 1017
Qy 241 AlaSerSerTyrSerLysAsnPheGlyLeuTyrAsnGluArgValGlyAlaCysThrLeu 260
Db 1018 GCCAGTCTCTACTCTAAAACCTTTGGCCTGTACAACGAGCGGTGTGGCGCTTGTACTCTG 1077
Qy 261 ValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaLeuArg 280
Db 1078 GTTGTGCGCAGCATGTAACCGTTGATCGCGCATTCAGCCAATGAAAGCGCGATTGCG 1137
Qy 281 AlaAsnTyrSerAsnProAlaHisGlyAlaSerValValAlaThrIleLeuSerAsn 300
Db 1138 GCTAACTACTCTAACCCACACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAAC 1197
Qy 301 AspAlaLeuArgAlaIleTrrGluGlnGluLeuThrAspMetArgGlnArgIleGlnArg 320
Db 1198 GATGGTTTACGTGCGATTGGAACAAGAGCTGATGATATGCGCAGCGATTATTCAGCGT 1257
Qy 321 MetArgGlnLeuPheValAsnThrLeuGlnLysGlyValAsnArgAspPheSerPhe 340
Db 1258 ATGCGTCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1317
Qy 341 IleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArg 360
Db 1318 ATCATCAACACAGACGCGATGTTCTCTTCACTGCGCTGACAAAGAACAAAGTGTGCGT 1377
Qy 361 LeuArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMet 380
Db 1378 CTGCGCAAGAGTTTGGCGTATATGCGTGTGCTTCTGCTGCGTAAATGTGCGCGGATG 1437
Qy 381 ThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 1438 ACACAGATATAACATGCTCCGCTGCGAAGCGATTGTGGCAGTGTGCTG 1485

```

RESULT 6

US-10-018-470B-1/c

```

; Sequence 1, Application US/10018470B
; GENERAL INFORMATION:
; APPLICANT: FRASER, Claire Marie
; APPLICANT: HICKEY, Erin Kathleen
; APPLICANT: PETERSON, Jeremy D.
; APPLICANT: TETTELIN, Herve
; APPLICANT: VENTER, J. Craig
; APPLICANT: MASIGNANI, Vega
; APPLICANT: GALEOTTI, Cesira
; APPLICANT: MORA, Marirosa
; APPLICANT: RATTI, Giulio
; APPLICANT: SCARSELLI, Maria
; APPLICANT: SCARLATO, Vincenzo
; APPLICANT: RAPPUOLI, Rino
; APPLICANT: PIZZA, Mariagrazia
; APPLICANT: GRANDI, Guido
; TITLE OF INVENTION: NEISSERIA GENOMIC SEQUENCES AND METHODS OF THEIR USE
; FILE REFERENCE: 223002100400
; CURRENT APPLICATION NUMBER: US/10/018,470B
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: PCT/US00/05928
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: US 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: GB 0004695.3
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/US99/23573
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: SeqWin99, version 1.0.4
; SEQ ID NO 1
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; US-10-018-470B-1

Alignment Scores:
Pred. No.: 2,9e-143 Length: 2242716
Score: 1310.00 Matches: 244
Percent Similarity: 78.0% Conservative: 64
Best Local Similarity: 61.8% Mismatches: 87
Query Match: 64.1% Indels: 0
DB: 7 Gaps: 0

US-10-673-786A-2 (1-396) x US-10-018-470B-1 (1-2242716)

Qy 2 PheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPheArg 21
Db 562972 TTCAAGCACATCGAAGCGCGCCCGCCGATTCGATTCTCGGTTTGGCGAAGCGTTCAAA 562913
Qy 22 AlaAspGluArgProGlyLysAlaLeuAsnLeuGlyLysValTyrLysAspGluThrGly 41
Db 562912 GCCGAAACCCCGCCGAAAGAGTCAACCTCGCGATCGCGGTTTATAAGACCGCATCCGCG 562853
Qy 42 LysThrProValLeuThrSerValLysLysAlaGluGlnTyrLeuLeuGluAsnGluThr 61
Db 562852 GCGACACCCCTCGTCAAGCGCGTCAAGAGCGCGAAGAAACCGCTGTGGAAAGCGAAACC 562793
Qy 62 ThrLysAsnTyrLeuGlyLysAlaLeuAspGlyLeuProGluPheGlyArgCysThrGlnGluLeu 81
Db 562792 ACCAAAACTACCTCACCATCGACGCGCTTCCGACTACACGCGCAAAACCCCAATCTCTG 562733
Qy 82 LeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGlnThrPro 101
Db 562732 CTGTTGCGCAAGACCAACCAATCATCCCGCGCGCGCCGCAAAACCGCAAGCGCTC 562673
Qy 102 GlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSerValLys 121
Db 562672 GCGCGTACGGCGCATTCGCTATTGCGCGCGGTTTGGCAACCGCAGTTGACCGCGCAA 562613
Qy 122 ArgValTrrValSerAsnProSerTrpProAsnHisLysSerValPheAsnSerAlaGly 141
Db 562612 ACCATCTGGATTTTCCAATCCGACTTGGCCCAACCAACGCGCATCGCCAAACGCGTCTG 562553

```

```

Qy 142 LeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPheAspAla 161
Db 562552 ATCCAAGACAAACCTTATCGTTACTATGATGCGCGCAACACGGTTTGGATTGGACGGC 562493

Qy 162 LeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGlyCysGys 181
Db 562492 ATGATTAGGAGCTTGGAGCAAGCGCAAAAGGCGACATCGTCTGCTCCACGGCTGCTGC 562433

Qy 182 HisAenProThrGlyIleAspProThrLeuGluGlnTnpGlnThrLeuAlaGlnLeuSer 201
Db 562432 CACAATCTACCGGATGACCTACGCGCGGAGCAACAAATGGAAACTTTTGGCAAACTTCT 562373

Qy 202 ValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArgGlyLeu 221
Db 562372 GCCGAAAGAGCTGCTGCGCGTGTTCACITTTGCTACCAAGGCTTCGGCAATGTTTG 562313

Qy 222 GluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeuIleValAla 241
Db 562312 GAAGAAGATGCTACGGCTTGGCGTGTTCCTTGAACAACAATACAGAATTCGATTGCC 562253

Qy 242 SerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyValaCysThrLeuVal 261
Db 562252 AGCTCTTATTCGAAACTTCGATATGTACACGAGCGCGTTCGCGCGCTTCACITTTGCTG 562193

Qy 262 AlaAlaAppSerGluThrValAspArgAlaPheSerGlnMetLysAlaAlaIleArgAla 281
Db 562192 GCCGAAGATGAAGAACACGACGCCGCCGCCACAGCCCAAGTCAAAACCATCATCGTACC 562133

Qy 282 AenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeuSerAenAsp 301
Db 562132 TTGTATTCCAAACCCGGCTTCACACGGTGCAGAACACCATTTGGCTGTTCGAAATGAT 562073

Qy 302 AlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgGlyLeuArgMet 321
Db 562072 GATTITGAAGACAAATGGATTGCCGAATTCGATGAAATTCGCGCGCGCATCAAGCCCATG 562013

Qy 322 ArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAspPheSerPheIle 341
Db 562012 CGCCAAAATTTGTCGGGTGCTCAAGCCCAAGGTGCAAGCCAAACTTTGATTTCATT 561953

Qy 342 IleLysGlnAenGlyMetPheSerPheSerGlyLeuThrLysGluGlnValLeuArgLeu 361
Db 561952 ATCAAAACAAACGGATGTTCTTTACGCGCTTGACTCCCGAACAAGTCGACCGCGCTG 561893

Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAenValAlaGlyMetThr 381
Db 561892 AAAAAACGAGTTTGCCTTTATGCGCTCCGCTCCGCGCGCGCATCAAGTCCGCGCATTACC 561833

Qy 382 ProAspAenMetAlaProLeuCysGluAlaIleValAlaValLeu 396
Db 561832 GACAACAACATCGATTATCTGTGCGAAGCATCGTGAAGTACTG 561788

```

RESULT 7

```

PCT-US05-47111-133
; Sequence 133, Application PC/TUS0547111
; GENERAL INFORMATION:
; APPLICANT: Abad Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: PCT/US05/47111
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 133
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Zea mays
PCT-US05-47111-133

```

```

Alignment Scores: 5.37e-100 Length: 1230
Pred. No.: 913.00 Matches: 180
Score:

```

```

Percent Similarity: 62.6% Conservative: 69
Best Local Similarity: 45.2% Mismatches: 147
Query Match: 44.6% Indels: 2
DB: 1 Gaps: 2

```

US-10-673-786A-2 (1-396) x PCT-US05-47111-133 (1-1230)

```

Qy 1 MetPheGluAenIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
Db 22 GTCTTCGCGCAGCTCGAGAGGAGCCCGAGGACCCCATCTCTCGAGGTGACCGTTGCCTAC 81

Qy 21 ArgAlaAspGluArgProGlyLysIleAenLeuGlyIleGlyValTyrLysAspGluThr 40
Db 82 AACAAAGATCCCAGCCCATGAAGGTCAACCTCGGGGTTCGCGCTACCCGACCGAGGAA 141

Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyrLeuLeuGluAenGlu 60
Db 142 GGGAAAGCCCTAGTGTGCTGAACGTGTCAGCGCGCCGAGCAAAATGTTGATCAATAATCCG 201

Qy 61 Thr---ThrLysAenTyrLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
Db 202 TCACGTGTCAAGGAGTAGTACCTACCAATCACCGGTCTGCTGAAATTCATAAAGTCAGCGCT 261

Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAenAspLysArgAlaArgThrAlaGln 99
Db 262 AAGCTTATCTTTGGCGTCGACGCCCTGCTATTCAAGAGAAATAGGGTTGTCTACCGTGCAG 321

Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAenThrSer 119
Db 322 TGCCTATCGGGTACTGTTCTTTAAGAGTCGGAGGTGAAATTTCTTGCAGGCACTATCAC 381

Qy 120 ValLysArgValTrpValSerAenProSerTrpProAenHisLysSerValPheAenSer 139
Db 382 GAGCGCATCTATCATCCACAAACCAACCTGGGGAATACCCCAAAAGTCTTCACCCCTA 441

Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTyrAspAlaGluAenHisThrLeuAspPhe 159
Db 442 TCTGGCTTGAACGTGTAGGAGTACCGCTATTATGATCCTGCAACATGCGAGCTTCACCTC 501

Qy 160 AspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
Db 502 GAAGGATCTATGGAAAGACCTCGGTTCTGCTCCTTCAGGTTCAATGTACTGTCATGCC 561

Qy 180 CysCysHisAenProThrGlyIleAspProThrLeuGluGlnTnpGlnThrLeuAlaGln 199
Db 562 TGTCCCAACACCTACTCGGAGTAGATCCTACCATCGAACACAGTGGGAACAGATTAGGCAG 621

Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyrGlnGlyPheAlaArg 219
Db 622 CTGATGAGATCAAAATCACTGCTTCCGTTCTTTCACAGTGCCTATCAAGGCTTTGCAAGT 681

Qy 220 Gly---LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
Db 682 GGAAGCTTTGACAAAGATGCTCAGTCAGTGGTATGTTTGTGCTGATGGTGGTGAACCTG 741

Qy 239 IleValAlaSerSerTyrSerLysAenPheGlyLeuTyrAenGluArgValGlyAlaCys 258
Db 742 CTCATGCTCAGAGCTACCGTAAGAACATGGATTGTATGGAGAGCGGTGTGGCGCTTTG 801

Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
Db 802 AGCATTTGTATGTAAGTGGCGATGTAGCTGTTAGGGTTGAAAGTCAACTCAAACTTCTGC 861

Qy 279 IleArgAlaAenTyrSerAenProAlaHisGlyAlaSerValValAlaThrIleLeu 298
Db 862 ATCAGGCTATGATTCAAAACCTCTCTTCATCGTGCCTCTATTGTTGTCTACCATCTC 921

Qy 299 SerAenAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgIle 318
Db 922 AGGACACGAGATGTTTCAACGAATGGACTCTGGAACCTGAAGGCGCATGGCTGATAGGATC 981

Qy 319 GlnArgMetArgGlnLeuPheValAenThrLeuGlnGlyLysGlyAlaAenArgAspPhe 338

```

Db 982 ATTAACATGAGGCAACAACCTATTATTAATGCGCTGAAATCCAGAGAAACCCCTGGTGATGG 1041
 Qy 339 SerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal 358
 Db 1042 AGCCATATCATTAAGCAAAATGGGATCTTTACTTTCTACCTGGCTGAATACGGAACAAGTG 1101
 Qy 359 LeuArgLeuArgGluGluPheGlyValTyraValaValaSerGlyArgValaAsnValala 378
 Db 1102 GCATTTCATGAGCGAGGAATACCACTATTATATGACATCTGATGGAGGATCAGCATGGCC 1161
 Qy 379 GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValaValLeu 396
 Db 1162 GGTTTGACATGAGGACTGTGCCCATCTTCAGATGCCATACACGTGCAGTT 1215
 RESULT 8
 US-60-752-355-133
 ; Sequence 133, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 133
 ; LENGTH: 1230
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-60-752-355-133
 Alignment Scores:
 Pred. No.: 5,37e-100 Length: 1230
 Score: 913.00 Matches: 180
 Percent Similarity: 62.6% Mismatches: 69
 Best Local Similarity: 45.2% Conservatives: 147
 Query Match: 44.6% Indels: 2
 DB: 12 Gaps: 2
 US-10-673-786A-2 (1-396) x US-60-752-355-133 (1-1230)
 Qy 1 MetPheGluAsnIleThrAlaAlaProAlaAspProIleLeuGlyLeuAlaAspLeuPhe 20
 Db 22 GTCTTCGCCGCATCTGAGCAGGCCCGGAGGACCCCATCTCTCGAGTGACCGTGGCTAC 81
 Qy 21 ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyrlsAspGluThr 40
 Db 82 AACAGGATCCAGCCCCCATGAAGGTCAACCTCGGGTTCGGCGCTACCGGACCGAGAA 141
 Qy 41 GlyIleThrProValLeuThrSerValLysAlaGluGlnTyrlsLeuGluAsnGlu 60
 Db 142 GGGGAAGCCCTTAGTGCTGAACGTGGTCAGCGCGCCGAGCAATGTTTGATCAATAATCCG 201
 Qy 61 Thr---ThrLysAsnTyrlsLeuGlyIleAspGlyIleProGluPheGlyArgCysThrGln 79
 Db 202 TCACGTCTCAAGGAGTACTACCAATCACCGGTCTGCTGAATTCATAAGCTGACGGCT 261
 Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAlaGln 99
 Db 262 AAGCTTATCTTTGGCGCTGACAGCCCTGCTATTTCAGGAGAAATAGGGTGTCTACCGTGCAG 321
 Qy 100 ThrProGlyThrGlyAlaLeuArgValaAlaAspPheLeuAlaLysAsnThrSer 119
 Db 322 TGCCTATCGGGTACTGTGTTCTTTAAGAGTCGAGGTGAATTTCTTCAAGAGCACTATCAC 381
 Qy 120 ValLysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
 Db 382 GAGCGCATCTATCATCCCAACCAACCTGGGGAATCACCCAAAGTCTTCCACCTA 441
 Qy 140 AlsGlyLeuGluValArgGluTyraIleTyrlsAspAlaGluAsnHisThrLeuAspPhe 159
 Db 442 TCTGGCTTGAACGTTAGAGTACCGGTATTATGATCCTGCAACATCGACGCTTCACTTC 501

```
US-10-673-786A-2 (1-396) x PCT-US05-47111-446 (1-1224)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 16 GTCTTCGCTCGCCCGCGAGCGCCGAGGACCCCATCTCGGGTGACGGTCGCGTAC 75
Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyLleGlyValTyrLysAaspGluThr 40
Db 76 AACAAAGATCCAGCCCGTGAAGTCAATCTCGCGCTCGCGGCTACCGGACCGAGAA 135
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluInTyrLeuLeuGluAenGlu 60
Db 136 GGAAGCCCTCGTGTGTAATGTGTCTGAGCGCGCTGAGCAGATGCTGATCAACAACCG 195
Qy 61 Thr---ThrLysAenTyrLeuGlyLleAaspGlyLeuProGluPheGlyArgCysThrGln 79
Db 196 TCAGTGTGAAGGAGTATTGGCCCAATCACTGGACTGGCCGATTTCAATAGCTGAGTGT 255
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGln 99
Db 256 AAGCTTATTTTGTGTGTCGACAGTCTCGCCATTCAGAGAAATAGGTGCTGACGTTACG 315
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAaspPheLeuAlaLysAenThrSer 119
Db 316 TGCTTGTGAGAACTGGTTCTTTAAGGTGGAGGTGAATTTCTTTCGAAGCATTTATCAT 375
Qy 120 ValLysArgValTyrValSerAenProSerTrpProAenHisLysSerValPheAenSer 139
Db 376 GAAGCGCATATATATATATATATATATATATATATATATATATATATATATATATAT 435
Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTrpAaspAlaGluAenHisThrLeuAaspPhe 159
Db 436 GCTGCGCTGACTGTAGGAGTTACCGTACTATGATCTCTGCAACCGCGGAGTGTGATT 495
Qy 160 AspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGly 179
Db 496 CAAGGCTGTTAGAAGATCTCGGTTTCAGCTCTCTGAGGTGCAATTTGTACTGCTTCT 555
Qy 180 CysCysHisAenProThrGlyLleAaspProThrLeuGluInTyrPheGlnThrLeuAlaGln 199

US-10-673-786A-2 (1-396) x US-60-752-355-446 (1-1224)
Qy 1 MetPheGluAenIleThrAlaAlaProAlaAaspProIleLeuGlyLeuAlaAaspLeuPhe 20
Db 16 GTCTTCGCTCGCCCGCGAGCGCCGAGGACCCCATCTCGGGTGACGGTCGCGTAC 75
Qy 21 ArgAlaAaspGluArgProGlyLysIleAenLeuGlyLleGlyValTyrLysAaspGluThr 40
Db 76 AACAAAGATCCAGCCCGTGAAGTCAATCTCGCGCTCGCGGCTACCGGACCGAGAA 135
Qy 41 GlyLysThrProValLeuThrSerValLysLysAlaGluInTyrLeuLeuGluAenGlu 60
Db 136 GGAAGCCCTCGTGTGTAATGTGTCTGAGCGCGCTGAGCAGATGCTGATCAACAACCG 195
Qy 61 Thr---ThrLysAenTyrLeuGlyLleAaspGlyLeuProGluPheGlyArgCysThrGln 79
Db 196 TCAGTGTGAAGGAGTATTGGCCCAATCACTGGACTGGCCGATTTCAATAGCTGAGTGT 255
Qy 80 GluLeuLeuPheGlyLysGlySerAlaLeuIleAenAaspLysArgAlaAargThrAlaGln 99
Db 256 AAGCTTATTTTGTGTGTCGACAGTCTCGCCATTCAGAGAAATAGGTGCTGACGTTACG 315
Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAaspPheLeuAlaLysAenThrSer 119
Db 316 TGCTTGTGAGAACTGGTTCTTTAAGGTGGAGGTGAATTTCTTTCGAAGCATTTATCAT 375
Qy 120 ValLysArgValTyrValSerAenProSerTrpProAenHisLysSerValPheAenSer 139
Db 376 GAAGCGCATATATATATATATATATATATATATATATATATATATATATATATATAT 435
Qy 140 AlaGlyLeuGluValArgGluTyrAlaTyrTrpAaspAlaGluAenHisThrLeuAaspPhe 159
Db 436 GCTGCGCTGACTGTAGGAGTTACCGTACTATGATCTCTGCAACCGCGGAGTGTGATT 495
Qy 160 AspAlaLeuIleAenSerLeuAenGluAlaGlnAlaGlyAaspValValLeuPheHisGly 179
Db 496 CAAGGCTGTTAGAAGATCTCGGTTTCAGCTCTCTGAGGTGCAATTTGTACTGCTTCT 555
Qy 180 CysCysHisAenProThrGlyLleAaspProThrLeuGluInTyrPheGlnThrLeuAlaGln 199

US-60-752-355-446
; Sequence 446 Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 446
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-752-355-446
Alignment Scores:
Pred. No.: 1,63e-99 Length: 1224
Score: 909.00 Matches: 179
Percent Similarity: 62.3% Conservative: 69
Best Local Similarity: 45.0% Mismatches: 148
Query Match: 44.4% Indels: 2
DB: 12 Gaps: 2
```


314	Db	GTCTTCAATCACTCTGGTTCTGCTCCCGAAGATCCTATCTCTCGGGGTAACTGTGCTGCTTAT	371
21	Qy	ArgAlaAspGluArgProGlyLysIleAsnLeuGlyIleGlyValTyLysAspGluThr	40
374	Db	AACAAAGATCCAACTCCAGTTAAGCTCAACTTGGGAGTTGGTCTTACCGAACTGAGAA	433
41	Qy	GlyLysThrProValLeuThrSerValLysLysAlaGluGlnTyLLeuLeuGluAsnGlu	60
434	Db	GGAAAACTCTGTTTGAATGTAGTGAGCGGAGTTGAACAGCAACTCATA--AATGAC	490
61	Qy	ThrThrLysAsn-----TyLLeuGlyLeAspGlyIleProGluPheGlyArgCysThr	78
491	Db	GTGTCCGCAACAAAGGAAATATATCCGATCGTTGGGCTTCTGATTTTAAATAATGAGT	550
79	Qy	GlnGluLeuLeuPheGlyLysGlySerAlaLeuIleAsnAspLysArgAlaArgThrAla	98
551	Db	GCTAAGCTTATTTTGGGGCTGACAGCCCTGCTATTTCAGACACACAGGTTACCACGTT	610
99	Qy	GlnThrProGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThr	118
611	Db	CAATGCTGTCTGGAACGTGTTCTTAAAGAGTTGGGGGTGAATTTTGGCTTAAACACTAT	670
119	Qy	SerValLysArgValTrpValSerAsnProSerTrpProHsnHisLysSerValPheAsn	138
671	Db	CACCAACGGGACTATATCTTCCCAACACCACTTGGGGCAATCACCAGAGTTTTCACAC	730
139	Qy	SerAlaGlyLeuGluValArgGluTyTyrAlaTyTyrAspAlaGluAsnHisThrLeuAsp	158
731	Db	TTTAGCGCTTGTCTGTCAAAACATACCGCTACTATGCTCCAGCAACACAGGACTTGAC	790
159	Qy	PheAspAlaLeuIleAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHis	178
791	Db	TTTTCAAGGACTTCTGGGAAGACCTTGGTTCTGCTCCATCTGGATCTATTGTTTCTCATAT	850
179	Qy	GlyCysCysHisAsnProThrGlyLysAspProThrLeuGluGlnTrpGlnThrLeuAla	198
851	Db	GCATGGCGCATAAACCCACTGTGTGGATGCCAACCCCTTGAGCAATCGGAGCAGATTAGG	910
199	Qy	GlnLeuSerValGluLysGlyTyTrpLeuProLeuPheAspPheAlaTyGlnGlyPheAla	218
911	Db	CAGCTAATAAGATCAAAAGCTTTGTTACCTTTCTTTGACAGTGCCTATCAGGGTTTGCT	970
219	Qy	ArgGly----LeuGluGluAspAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGlu	237
971	Db	AGTGGGAAGTCTAGATGCAGATGCCAACCTGTTCTGTTGTTGTTGCTGAGGAGCGAA	1030
238	Qy	LeuIleValAlaSerSerTySerLysAsnPheGlyLeuTyRanGluArgValGlyAla	257
1031	Db	TTGCTGTGTAGCAAAAGCTATGCAAAAGATCTGGGTCTTTATGGGAAACGTGTTGGCGCC	1090
258	Qy	CysThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAla	277
1091	Db	TTAAGCATTTCTGCAAGTCAAGTGTGATGTTGCAAGCAGGGTTGAGAGCCAGCTGAGACTA	1150
278	Qy	AlaIleArgAlaAsnTySerAsnProAlaHisGlyAlaSerValValAlaThrIle	297
1151	Db	GTGATTAGGCCCATGTACTCAAGTCTCCCATTCATGTTGATGCCATCCATTGTCGCTGCCATT	1210
298	Qy	LeuSerAsnAspAlaIleArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArg	317
1211	Db	CTCAAGACGCGGAATTTGTTTCAATGACTGCACTGATATTGAGTTGAAGGCAATGGCTGATCGC	1270
318	Qy	IleGlnArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyAlaAsnArgAsp	337
1271	Db	ATCATCAGTATGCGCAAGAACTTTTCGATGCTTTTATGTTCCAGAGGCCACACTGGCGCAT	1330
338	Qy	PheSerPheIleIleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGln	357
1331	Db	TGGAGTCACATATCAACACAGATTGGAATGTTTACTTCTCGATTGAATGCGGAACAA	1390
358	Qy	ValLeuArgLeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnVal	377
1391	Db	GTGTTCTTCACTAGTAAAGAGTTCCATATACATGACATCTCATGCGAGGATTAGCATG	1450

Percent Similarity: 59.0% Conservative: 73
 Best Local Similarity: 40.7% Mismatches: 161
 Query Match: 39.8% Indels: 2
 DB: 10 Gaps: 2

US-10-673-786A-2 (1-396) x US-11-299-286-563 (1-1218)

Qy 1 MetPheGluAenileThrAlaAlaProAlaAspProileLeuGlyLeuAlaAspLeuPhe 20
 Db 10 ATCTTGTCAAGCGCTCTTCTGCTCCCGAAGATCTGTTCTCAGTGAATTTTCTTGT 69
 Qy 21 ArgAlaAspGluArgProGlyLysleAsnLeuGlylleGlyValTyLysAspGluThr 40
 Db 70 AGAGATGATCTAGTCCGCTTAAGTTGAATTTGAGCGCAGGTACCTATCGAACTGAGGAA 129
 Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyLysLeuGluAen-- 59
 Db 130 GGAAGCCTCTTGTCTTGTGATGTTGTGAGAGAGCAGACACAGTTAGCCACACCTG 189
 Qy 60 GluThrThrLysAsnTyLeuGlylleAspGlylleProGluPheGlyArgCysThrGln 79
 Db 190 TCTAGGACAGGAATATCTCCCTCAATGACCTCTGTAATTTAATAAATTAAGTACC 249
 Qy 80 GluLeuLeuPheGlyLysSerAlaLeuLeuLeuAsnAspLysArgAlaArgThrAlaGln 99
 Db 250 AAGCTCATCTTAGGTGATGATCTCTGCTGATTTGAAAGAGAAATAGAGTTGTTACAAACCCAA 309
 Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
 Db 310 TGTGTTGCTGTGACGTGTTCTTGTGAGAGTTGAGCTGATTTCTAGCAACACACAA 369
 Qy 120 VallysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
 Db 370 GAAAGTGTCTTCTGTTCCAAACCCACAGTGGGGGAACCATCCCGGATTTTACTTGT 429
 Qy 140 AlaGlyLeuGluValArgGluTyAlaTyTyAspAlaGluAsnHisThrLeuAspPhe 159
 Db 430 GCGGTTTGTCTGTCAGTATTTCCGTATTTATGATCGGAGCGGAGGACTCGACTTC 489
 Qy 160 AspAlaLeuLeuAsnSerLeuAsnGluAlaGlnAlaGlyAspValValLeuPheHisGly 179
 Db 490 AAAGCATGCTGAGGATCTTGGCGCTGCACCTCTGAGCTTATAGTGTGCTCAAGCT 549
 Qy 180 CysCysHisAsnProThrGlyLeuAspProThrLeuGluGlnTrpGlnThrLeuAlaGln 199
 Db 550 TGTGCGCATACCCACAGGATTTGACCCCAATTCGAACATGCGAAAGATTCGACGA 609
 Qy 200 LeuSerValGluLysGlyTrpLeuProLeuPheAspPheAlaTyTrpGlnGlyPheAlaArg 219
 Db 610 CTAGTGAGATCGAAAGCTTATTACCTTCTTTGATGATGTCATATCAGGTTTGTCTAGT 669
 Qy 220 Gly----LeuGluLysAlaGluGlyLeuArgAlaPheAlaAlaMetHisLysGluLeu 238
 Db 670 GGTAGCTTGGCGAGATGCAACGCTTCTGATTTGTTAGCCGATGGAGGTGAATGT 729
 Qy 239 lleValAlaSerSerTySerLysAsnPheGlyLeuTyAsnGluArgValGlyAlaCys 258
 Db 730 TTGATGCTCAAGTTATGCCAANAATATGGGCCCTTTATGGGAGCGGATTTGGCTCTCT 789
 Qy 259 ThrLeuValAlaAlaAspSerGluThrValAspArgAlaPheSerGlnMetLysAlaAla 278
 Db 790 ACGATGATGATCATCAGAGATGCTCGTAAGAAAGTTGAGAACCAAGTCTGTTGT 849
 Qy 279 lleArgAlaAsnTySerAsnProAlaHisGlyAlaSerValValAlaThrIleLeu 298
 Db 850 GTGAGGCTATGATCTTACACACCTATTCATGTTGATGCTATTTGTTGCCAATTTCTA 909
 Qy 299 SerAsnAspAlaLeuArgAlaIleTrpGluGlnGluLeuThrAspMetArgGlnArgile 318
 Db 910 AAAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
 Qy 319 GlnArgMetArgGlnLeuPheValAsnThrLeuGlnGluLysGlyValAlaAsnArgAspPhe 338

Db 970 ATTAGCATGCTCAACAGTTATATATGACCTCTAGAGCTAGAGGACACCTGCTGATTGG 1029
 Qy 339 SerPheIlelleLysGlnAsnGlyMetPheSerPheSerGlyLeuThrLysGluGlnVal 358
 Db 1030 AGTCACATTATTAACATATTTGGGATGTTTACTTTTACCGGATTTGAGTGAGGCAAGTT 1089
 Qy 359 LeuArgLeuArgGluGluPheGlyValTyAlaValAlaSerGlyArgValAsnValAla 378
 Db 1090 CGCTTGTATGGCCAAAGAAATATACATTATACATGCTTATGATGGGAGGATAGCATGGCA 1149
 Qy 379 GlyMetThrProAspAsnMetAlaProLeuCysGluAlaIleValAlaValLeu 396
 Db 1150 AGTCTAAGTTCGAAGACAGTGCCTCACTGCTGATGCTATACATGCTGTTGTT 1203

RESULT 14

US-10-953-349-11025
 ; Sequence 11025, Application US/10953349
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: ENCODED THERBY
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 11025
 ; LENGTH: 1301
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-953-349-11025

Alignment Scores:
 Pred. No.: 5,698-87 Length: 1301
 Score: 806.00 Matches: 161
 Percent Similarity: 58.8% Conservative: 73
 Best Local Similarity: 40.5% Mismatches: 162
 Query Match: 39.4% Indels: 2
 Gaps: 2
 DB:
 US-10-673-786A-2 (1-396) x US-10-953-349-11025 (1-1301)

Qy 1 MetPheGluAenileThrAlaAlaProAlaAspProileLeuGlyLeuAlaAspLeuPhe 20
 Db 17 ATCTTGTCAAGCGCTCTTCTGCTCCCGAAGATCTGTTCTCAGTGAATTTTCTTGT 76
 Qy 21 ArgAlaAspGluArgProGlyLysleAsnLeuGlylleGlyValTyLysAspGluThr 40
 Db 77 AGAGATGATCTAGTCCGCTTAAGTTGAATTTGAGCGCAGGTACCTATCGAACTGAGGAA 136
 Qy 41 GlyLysThrProValLeuThrSerValLysAlaGluGlnTyLysLeuGluAen-- 59
 Db 137 GGAAGCCTCTTGTCTTGTGATGTTGTGAGAGAGCAGACACAGTTAGCCACACCTG 196
 Qy 60 GluThrThrLysAsnTyLeuGlylleAspGlylleProGluPheGlyArgCysThrGln 79
 Db 197 TCTAGGACAGGAATATCTCCCTCAATGACCTCTGTAATTTAATAAATTAAGTACC 256
 Qy 80 GluLeuLeuPheGlyLysSerAlaLeuLeuLeuAsnAspLysArgAlaArgThrAlaGln 99
 Db 257 AAGCTCATCTTAGGTGATGATGCTCTGCTGATTTGAAAGAGAAATAGAGTTGTTTACACCCAA 316
 Qy 100 ThrProGlyGlyThrGlyAlaLeuArgValAlaAlaAspPheLeuAlaLysAsnThrSer 119
 Db 317 TGTGTTGCTGTGACGTGTTCTTGTGAGAGTTGAGCTGATTTCTAGCAGCAGCACAAACAA 376
 Qy 120 VallysArgValTrpValSerAsnProSerTrpProAsnHisLysSerValPheAsnSer 139
 Db 377 GAAATGCTCATTTCTGTTCCAAACCCAACTTGGGGGAACCATCCCGGATTTTACTTGT 436
 Qy 140 AlaGlyLeuGluValArgGluTyAlaTyTyAspAlaGluAsnHisThrLeuAspPhe 159
 Db 437 GCGGTTTGTCTGTCAGTATTTCCGTATTTATGATCGGAGCGGAGGACTCGATTTC 496

Db 1131 ACCGACCAATGGCATGTTCTGTTTCACAGGGCTAAAGCCTGAACAGGTGGAGCGGCTG 1190
 Qy 362 ArgGluGluPheGlyValTyrAlaValAlaSerGlyArgValAsnValAlaGlyMetThr 381
 Db 1191 ATCAAGGAGTTCTCCATCTACATGACAAAAGATGGCGCATCTCTGTGGCAGGGGTCACC 1250
 Qy 382 ProAspAsnMetAlaProLeuCysGluAlaIle 392
 Db 1251 TCCAGCAACGTGGGCTACCTTGCCCATGCCATT 1283

Search completed: March 15, 2006, 16:06:44
 Job time : 1265 secs

This Page Blank (uspto)

Run on: March 15, 2006, 14:30:49 ; Search time 1436 Seconds
(without alignments)
1933.887 Million cell up

Gapop 10.0 , Gapext 1.0

Maximum DB seq length: 200000000

Database : Published Applications NA New:*

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1176.6	98.8	1191	12	US-11-114-922-75	Sequence 75, Appl
2	461.4	38.7	1269	8	US-10-467-657-6535	Sequence 6535, Ap
3	285.4	24.0	1191	8	US-10-467-657-1059	Sequence 1059, Ap
4	255.4	21.4	1194	8	US-10-979-821-31	Sequence 31, Appl
5	255.4	21.4	1194	12	US-11-114-922-31	Sequence 31, Appl
6	241.6	20.3	1260	8	US-10-979-821-5	Sequence 5, Appli
7	241.6	20.3	1260	12	US-11-114-922-5	Sequence 5, Appli
8	240	20.2	1260	8	US-10-979-821-3	Sequence 3, Appli
9	240	20.2	1260	12	US-11-114-922-3	Sequence 3, Appli
10	221.6	18.6	2343	12	US-11-136-527-2278	Sequence 2278, Ap
11	206.2	17.3	1170	8	US-10-979-821-1	Sequence 1, Appli
12	206.2	17.3	1170	12	US-11-114-922-1	Sequence 1, Appli
13	145.8	12.2	1239	8	US-10-979-821-7	Sequence 7, Appli
14	145.8	12.2	1239	12	US-11-114-922-7	Sequence 7, Appli
15	145.2	12.2	1251	7	US-10-932-182A-82350	Sequence 82350, A
16	145.2	12.2	1251	7	US-10-932-182A-82350	Sequence 82350, A
17	127.6	10.7	1389	7	US-10-932-182A-4671	Sequence 4671, Ap
18	127.6	10.7	1389	7	US-10-932-182A-4671	Sequence 4671, Ap
19	102	8.6	1400	12	US-11-136-527-6374	Sequence 6374, Ap
20	85	7.1	560	6	US-09-935-065A-479173	Sequence 479173, A


```

; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979, 821
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 31
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-10-979-821-31

Query Match      21.4%; Score 255.4; DB 8; Length 1194;
Best Local Similarity 51.5%; Pred. No. 9.2e-73;
Matches 614; Conservative 0; Mismatches 576; Indels 3; Gaps 1;

QY 2 TGTGAGAACATTTACCGCGCTCTCGCGACCCGATTTCTGGCCCTGGCCGATCTGTTTC 61
DB 2 TGTTCACAAAGTTGACGCTACGCTGGCGACCGGATTTCTTACGCTTATGGAGCGTTTAA 61

QY 62 GTGCCGATGAAGCTCCCGGCAAAATTAACCTCGGATTTGTGCTATATAGATGACAGCG 121
DB 62 AAGAAGACCCCTCGCAGCGACAAAGTGAATTTAAGTATCGGTCTGTACTACAAACGAAG 121

QY 122 GCAAAACCCCGTACTACCGAGCGTGAAGAGCTGAACAGTATCTGCTCGAAAA---TG 178
DB 122 GAATATTCCAAACTCAAGCGCTGGCGAGGCGGAGCGCGCTGAATGCGAGCTC 181

QY 179 AAACCCACCAAAATTTACTCGGCATTTGACGGCATCCCTGAAATTTGGTTCGTGCACTCAGG 238
DB 182 ATGGCGCTCGCTTTATTACCGATGAAGGGCTTAACCTGCTATCGCATGTCATTCGCG 241

QY 239 AACTGCTGTTTGGTAAAGTAGCCCTGATCAATGACAAACGTCGTCGACGCGACAGA 298
DB 242 CGCTGCTGTTTGGTGGCGACCATCCGGTACTGAAACAAACAGCGGTAGCAACCAATCAA 301

QY 299 CTCGGGGGCACTGGCGCACTACGCGTGTGCTCGGATTTCTGGCAAAAAATACACAGCG 358
DB 302 CCCTTGGCGGCTCGGGGCAATTGAAAGTGGCGGGATTTCTTGAAACGCTACTTCCCGG 361

QY 359 TTAAGCGTGTGTTGGTGAGCAACCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTG 418
DB 362 AATCAGCGCTCTGGTCAGCGATCTACCTGGGAAACCGGTAGCAATATTCGCGGG 421

QY 419 CAGTCTGGAAGTTGCTGAATACGCTTATATGATGCGGAAATCAACACTCTTGACTTCG 478
DB 422 CTGGATTCGAAGTAGTACTTACCCCTGGTATGACGAAGCGACTTAACGGCGTGGCGTTA 481

QY 479 ATGACACTGAATTAACAGCGCTGAATGAAGCTCAGGCTGGCGAGTGTCTGTTCCATGGCT 538
DB 482 ATGACCTGTTGGCGACGCTGAACAACTTACTGCCCCGAGTATTTGTTGCTGCAATCCAT 541

QY 539 GCTGCCATAACCCAAACCGGTATGACCCCTACGCTGGAAACAATGGCAAACTGCGCAAC 598
DB 542 GTTGCCACAAACCAACGGGTGCCGATCTCAATATGATCAGTGGATGCGGTGATTGAAA 601

QY 599 TCTCCGTTGAGAAAGGTGTTTACCGCTGTTTGAATCTTCGCTTACCAAGGTTTTCGCCGTG 658
DB 602 TTCTCAAAAGCCCGGAGCTTATTCATTCCTCGATATTTGCTATCAAGGATTTGGTCCG 661

QY 659 GTCTGGAAGAGATGCTGAAGGACTGCGCGCTTTTCGGCGCTATGCAATAAGAGCTGATTG 718
DB 662 GTATGGAAGAGATGCTGCTATTCGCGCAATTGCGAGCGTGGATTTACCGCTCTGG 721

```

RESULT 5

```

US-11-114-922-31
; Sequence 31, Application US/111114922
; Publication No. US2005028260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; PRIOR FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 31
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-11-114-922-31

```

```

Query Match      21.4%; Score 255.4; DB 12; Length 1194;
Best Local Similarity 51.5%; Pred. No. 9.2e-73;
Matches 614; Conservative 0; Mismatches 576; Indels 3; Gaps 1;

QY 2 TGTGAGAACATTTACCGCGCTCTGCGACCCGATTTCTGGCCCTGGCCGATCTGTTTC 61
DB 2 TGTTCACAAAGTTGACGCTACGCTGGCGACCCGATTTCTTACGCTTATGGAGCGTTTAA 61

QY 62 GTGCCGATGAAGCTCCCGGCAAAATTAACCTCGGATTTGTGCTATATAGATGACAGCG 121
DB 62 AAGAAGACCCCTCGCAGCGACAAAGTGAATTTAAGTATCGGTCTGTACTACAAACGAAG 121

QY 122 GCAAAACCCCGTACTACCGAGCGTGAAGAGCTGAACAGTATCTGCTCGAAAA---TG 178
DB 122 GAATATTCCAAACTGCAAGCGCTGGCGAGGCGGAGCGCGCTGAATGCGAGCTC 181

```

179 AAACACCAAAATTAACCTCGGCATTAACCGCATCCCTGAATTTGGTGGCTGCACTCAGG 238
 182 ATGGCGCTCGCTTTATTACCGATGGAAGGCTTAACCTGCTATCGCATGCTCAGG 241
 239 AACTGCTGTTGGTAAAGTAGAGCCCTGATCAATAGCAAAAGCTGCTCGACGGCAAGA 298
 242 CGCTGCTGTTGGTGGCGACCATCCGCTACTGAAACCAACAGCGGCTAGCAACCAATCAAA 301
 299 CTCGGGGGACCTGGCGCACTACGCTGGCTGCGGATTTCTCGCAAAATAACACGCG 358
 302 CCCTTGGCGCTCGGGGATTTGAAAGTGGCGGCGATTTCTGAAACGCTACTTCCCGG 361
 359 TTAAGCGTGTGGGTGAGCAACCAAGCTGGCGCAACCAATAAGAGGCTCTTTAACTCTG 418
 362 AATCAGCGCTCGGTCAGCGATCTACCTGGGAACCGCGTAGCAATATTCGCCGGG 421
 419 CAGGCTGGGAAGTTGCGTAATACGCTTATTAATGATGCGGAAATCAACACTTTCGACTTCG 478
 422 CTGGATTCGAAGTGAGTACTTACCCCTGGTATGAGAAAGGAGCTTAACGGGCTGCGCTTTA 481
 479 ATGCACATGATTAAACAGCTGAATGAAGCTCAGGCTGGCGAGTAGTCTGTTCATCGCT 538
 482 ATGACCTGTGGCGACGCTGAAACATTAACCTGCCCGAGTATTGTGTGTCATCCAT 541
 539 GCTGCCATAACCAACCGGTATGACCTCTAGCTGGAAACAATGGCAACACTGGCAAC 598
 542 GTTGCCACCAACCAACCGGTGCCATCTCACTAATGATCACTGGATGCGGTGATTGAAA 601
 599 TCTCCGTTGAGAAAGGTGTTACCGCTGTTGATCTGCTTACCGGTTTACCGGTTTCCCGTG 658
 602 TTCTCAAGCCCGGAGCTTATTCATTCCTCGATATTGCTATCAAGGATTTGGTCCG 661
 659 GTCTGGAAGAGAGCTGAAGACTGCGGCTTTCGCGGCTATGCTAATAAGAGCTGATTG 718
 662 GTATGGAAGAGGATGCTTACGCTATTTCGCGCAATTGCGAGCTGGAATTACCGCTCTGG 721
 719 TTGCGAGTCTCTACTCTAAAACCTTTGGCCCTGTACAACGAGCGTGTGGCGCTTGTACTC 778
 722 TGACCAATTCGTTCTGANAATTTTCCTCTTACGGGAGCGGCTCGGCGGACTTTCG 781
 779 TGGTGTGCTCGCAAGTAAACCGTTGATCGCGATTCAGCAAAATGAAAGCGCGGATTC 838
 782 TTATGTGTGAAGATGCGAAGCCGCTGGCGGCTACTGGGCAATTTGAAAGCAACAGTTC 841
 839 CGGCTAATCTACTTAACCCACCAAGCAACGCGGCTTCTGTTGTCACCATCTCTGACA 898
 842 GCGCAACTACTCCAGCCCGCGGAATTTTGGTGGCGAGGTGGTGGCTGCAAGTCTGAATG 901
 899 ACGATGGTTACGTGCGAATTTGGGAACAAGAGCTGACTGATATGCGCGCAGCGTATTCAGC 958
 902 ACGAGGCAATTAAGAGCCAGCTGGCTGGGAGAGTAGAAGAGATGCGTACTCGCATCTCG 961
 959 GTATGCGTCAAGTGTTCGTAATACGCTGAGGAAAAAGGCGCAAAACCGGACTTTCAGCT 1018
 962 CAATGCGTCAAGAAATGTTGAAGTATTAAAGCAGAGATGCCAGAACGCAATTTTCGATT 1021
 1019 TTATCATCAACAGAACCGGATGTTCTCTTCAGTGGCCTTGACAAAAGAACAGTGTCTGC 1078
 1022 ATCTGCTTAATACGCGGCGATGTTCAAGTTATACCGGTTTAAAGTGGCGCTCAGGTTGACC 1081
 1079 GTCTGCGCAAGAGTTTGGGCTATATCGGTTGCTTCTGCTGGGTAAATGTTGCGCGGA 1138
 1082 GACTACGTGAAGAAATTTGGTGTCTATCTCATCGCAGCGGTGCGATGTGTGCGCGGGT 1141
 1139 TGACACAGATTAACATGGCTCCGCTGTGCGAAGCGATTGTGGCAGTCTGTAA 1191
 1142 TAAATACGCAAAATGTACAAGTGTGGCAAGGCGTTTGTGCGGTGATGTAA 1194

Publication No. US20050244937A1
 GENERAL INFORMATION:
 APPLICANT: ABRAHAM, TIMOTHY W.
 APPLICANT: CAMERON, DOUGLAS C.
 APPLICANT: HICKS, PAULA M.
 APPLICANT: MCFARLAN, SARA C.
 APPLICANT: MILLIS, JIM
 APPLICANT: ROSAZZA, JACK
 APPLICANT: ZHAO, LISHAN
 APPLICANT: WEINER, DAVID P.
 TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
 PRODUCTION OF MONATIN AND ITS PRECURSORS
 FILE REFERENCE: 023829-0390
 CURRENT APPLICATION NUMBER: US/10/979,821
 CURRENT FILING DATE: 2004-11-03
 PRIOR APPLICATION NUMBER: 10/422,366
 PRIOR FILING DATE: 2003-04-23
 PRIOR APPLICATION NUMBER: 60/374,831
 PRIOR FILING DATE: 2002-04-23
 NUMBER OF SEQ ID NOS: 77
 SOFTWARE: PatentIn Ver. 3.3
 SEQ ID NO 5
 LENGTH: 1260
 TYPE: DNA
 ORGANISM: Rhodobacter sphaeroides
 US-10-979-821-5

Query Match 20.3%; Score 241.6; DB 8; Length 1260;
 Best Local Similarity 51.8%; Pred. No. 3.1e-68;
 Matches 603; Conservative 0; Mismatches 549; Indels 12; Gaps 2;

QY 25 CCTGCCGACCCGATTTCTGGCGCTTGGCGATCTGTTTCGTCGATGAACGTCGCGAAA 84
 DB 100 CCGCGGACAAGATCTGCAACTGATCCAGATGTTCCGCGAGGATGCGCGCGGACAG 159
 QY 85 ATTAACCTCGGATTTGGTGTCTATAAAGATGAGACGGGCAAAACCCCGGTACTGACGAGC 144
 DB 160 ATCGATCTGGCGCTGGCGCTTACAAGGACCCGACCGGGCTCACCCGGTCTATCGGGCC 219
 QY 145 GTGAAAAGGCTGAACAGTATCTGTCGAAAATGAACACCAACCAAAATTAACCTGGCATTT 204
 DB 220 GTGAAGCCCGCAGAAAGCGGCTCTGGGAGGTGAGAGCCACCAAGACTACACCGGCTT 279
 QY 205 GACGGATCCCTGAATTTGGTGTGCTGCACTCAGGAATGCTGTTTGGTAAAGTAGGCGCC 264
 DB 280 GCGCGGAGCCCGCTACATGCGCGATGCGGAGCTGAGTCTCTCG-----CAGGC 330
 QY 265 CTGATCAATGACAAAAGTGTCTCGACGCGACAGATCCGGGGGCGCATGCGCGCACTACGC 324
 DB 331 CGGTTCCCGCCGACCGGGTGGCTCGGTCCGACACCCCGCGGCGACGCGCGGTGCGT 390
 QY 325 GTGGCTGCCGATTTCTGGCAAAAATACAGCGTTAAGCGTTAAGCGTGTGGTGAAGCAACCA 384
 DB 391 CAGGCGCTCAGAGTGTATCCGCATGCGCTCCCGAGGCCACTGTGCTGGAATCTCGAACCCG 450
 QY 385 AGCTGGCGCAACCATTAAGAGCGCTTTTAACTCTGCGAGTCTGGAAGTTCTGTAATACGCT 444
 DB 451 ACCTGGCGCAACCATCTGTGATGCGTAATATCTCGGCATCCCGATCCCGGATATACCG 510
 QY 445 TATTATGATGCGGAAAATCACTCTTTGACTTCGATGCACTGATTAACAGCCTGAATGAA 504
 DB 511 TATTTCGACCCGAGACCGCGCGCTGATGCGAGGGCTTGTATGGAGGATCTGCGCCAG 570
 QY 505 GCTCAGCTGGCGACGATGCTGTTCCATGCGTGTGCGCATACCAACCAACCGGTATCGAC 564
 DB 571 GTGAAGCGGCGACGCTGCTGCTGACCGGTGTGCGCAACACCCGACCGCGCGCAAC 630
 QY 565 CCTACGCTGGAACAATGGAACAAACACTCTCCGTTGAGAAAGGCTGCTGTTACCG 624
 DB 631 CCGAACCCGTTGAGTGGCTGGCGCTCTGCGAGAGCTGCGCCGACAGCGCGGTGCGG 690
 QY 625 CTGTTTGAATTCGCTTACCAAGGGTTTTCGCCGTTTTCGGAAGAGATGCTGAAGAGATG 684

Db 691 CTGATCGACCTCGCCTATCAGGGCTTCGGCGACGGGCTCGAGATGGATCGGCGCGCAGC 750
 Qy 685 CGCGCTTCGCGGCTATGCAATAAGAGCTGATTTGTTGCCAGTTCTCTACTCTAAAACTTTT 744
 Db 751 CGGCTTCGCGCACCAGACTCGCCGAGGTCTGATCGCGGCTCTCTCTCGAAGAACTTC 810
 Qy 745 GGCTGTACAAACAGCGGTGTGGCGCTTGTTACTCTGTTGCTCGCAGCAGTGAACCGTT 804
 Db 811 GGCATCTACCGCGAGCGAACCGGCATCTCTGA---TCGCCATCGCGCGAGCGCGCGCG 867
 Qy 805 GATCGGCGATTCAGCGCAATGAAAGCGGCGATTCGGCTAACTACTCTAACCCACAGCA 864
 Db 868 GGCACGGTGCAGGCCAACCTCACTTCCTGAAACGGCGAGAACTACTCTCTCCCGCGGAC 927
 Qy 865 CACGGCGCTTCTGTTGTGCCACCATCTCTGAGCAACGATGCGTTACGTCGATTTGGGAA 924
 Db 928 CATGGCGCGGCTCGTGACCATGATCTCTGAGGACGAGACGCTGAGCGCGGACTGGAAG 987
 Qy 925 CAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGTATGCGTCAGTTGTTCTGATACG 984
 Db 988 GCGGAACTCGAGGAGGTGCGGCTCAACATGCTGACGCTGCGCGCGCAGCTTCGCGATCG 1047
 Qy 985 CTCAGGAAAGGCGCAACCGCACTTCAGCTTTATCATCAAAACAGAACGCGATGTTT 1044
 Db 1048 CTGAGGCCGAGACCGGCTCGAACCGCTTCGGCTTCGTTGCGCGCAGCATTCGCGCATG 1107
 Qy 1045 TCCTTCAGTGGCTGACAAAGAACCAAGTCTCGTTCGCGAAGAGTTTGGCGTATAT 1104
 Db 1108 TCGCGCTCGGATCAGCCCGCGAGGTGGAGCGGCTCGGACCGAGCAGCGGGTCTAC 1167
 Qy 1105 GCGGTTGCTTCTGTTGCGGTAATGTTGGCCGGGATGACACAGATACATGCGTCCGCTG 1164
 Db 1168 ATGGTGGCGATTCGCGGCTGAAACATCGCGGGCTGAAACCGGAGCGACCGTGCCTG 1227
 Qy 1165 TGCAGAGCGATTGTGGCAGTGTCTG 1188
 Db 1228 GCGCGCGGTGCGCAAGGTGCTG 1251

RESULT 7
 US-11-114-922-5
 ; Sequence 5, Application US/11114922
 ; Publication No. US2005028260A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HICKS, PAULA M.
 ; APPLICANT: MCFARLAN, SARA C.
 ; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
 ; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
 ; FILE REFERENCE: 023829-0396
 ; CURRENT APPLICATION NUMBER: US/11/114,922
 ; PRIOR FILING DATE: 2005-04-26
 ; PRIOR FILING DATE: 10/422,366
 ; PRIOR FILING DATE: 2003-04-23
 ; PRIOR APPLICATION NUMBER: 60/374,831
 ; PRIOR FILING DATE: 2002-04-23
 ; NUMBER OF SEQ ID NOS: 91
 ; SOFTWARE: PatentIn Ver. 3.3
 ; SEQ ID NO 5
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Rhodobacter sphaeroides
 ; US-11-114-922-5

Query Match 20.3%; Score 241.6; DB 12; Length 1260;
 Best Local Similarity 51.8%; Pred. No. 3.1e-68;
 Matches 603; Conservative 0; Mismatches 549; Indels 12; Gaps 2;
 Qy 25 CTTGCCGACCCGATTTCTGGCGCTGGCGATCTGTTCTGTCGCCGATGAACCTCCGCGAAA 84
 Db 100 CCGCGGACAAAGATCTCTGCAACTGATCCAGATGTTCCGCGAGGATGCGCGCGGACAAG 159
 Qy 85 ATTAACCTCGGATTTGTTGTTATTAAGATGAGACGGCGGCAAAACCCCGGTACTGACGAC 144

Db 160 ATCGATCTGGCGTGGCGCTTACAAGGACCCGACCGGGCTCACCCCGGCTCATGCGGGCC 219
 Qy 145 GTGAAAAAGGCTGAACAGATATCTGCTCGAAAAATGAACACCACCAAAAATTTACCTCGCAT 204
 Db 220 GTGAAGGCGCGCGAGAGCGGCTCTGGAGGTTCGAGACCAACCAAGACCTTACACCGGCTT 279
 Qy 205 GACGCGATCCCTGAAATTTGGTTCGCTGCACTCAGGAACTGCTGTTTGGTAAAGGTAGCGCC 264
 Db 280 GCGCGGAGCCCGCTTCAATGCCGGGATGGCGAAGCTGATCTCTCG-----CAGGC 330
 Qy 265 CTGATCAATGACAAAACGTCCTCGCAGCGCACAGACTCCGGGGGGGCACTGCGGCATACGC 324
 Db 331 GCGGTCGCGCGCGACCGGGTGGCTCGTTCGCGCACCCCGCGGCGCACCGGCGCGGTGCGT 390
 Qy 325 GTGGGTGCGCATTTCTTGGCAAAAATACAGGTTTAAGCGTGTGTTGGGTGAGCAACCCA 384
 Db 391 CAGGCGCTCGAGCTGATCCGATGGCTTCGCCGAGGCCACTGTCTGGAATCTCGAACCCG 450
 Qy 385 AGTGGCCGAAACATAAGAGCGTCTTTAACTCTGCAAGGTCTTGGAACTTCGTGAATACGCT 444
 Db 451 ACTGCCCCGAAACATCTGTCGATTCGTAATATCTCGGCATCCCGATGGCGAATACCGC 510
 Qy 445 TATTATGATCGGAAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCTGAATGAA 504
 Db 511 TATTTCGACGCCGAGACCGGCGCGCTCGATGCGAGGGCTTGATGAGGATCTGGCCAG 570
 Qy 505 GCTCAGGCTGGGCAAGTGTGTCATGGTGTCTGCATTAACCAACCAACCGGTATTCGAC 564
 Db 571 GTGAAGGCGGGGCAAGTGTGTCGTCACGGCTGCTGCACAAACCGGCGCGCAAC 630
 Qy 565 CTTACGCTGGAACAAATGGCAAAACACTGGCACAACTCTCGTTTGAGAAAGGCTGTTTACCG 624
 Db 631 CGGAACCGGTGAGTGGCTGCGCGTCTGCGAGAGCTGCGCGGACAGCGCGGTGCGCG 690
 Qy 625 CTGTTTGACTTCGCTTACAGGGTTTTCGCCGTGTTGCGAAGAAAGATGCTGAAGACTG 684
 Db 691 CTGATCGACTTCGCTTATCAGGGCTTCGCGCAGGGCTCGAGATGATGCGGCGCGGACG 750
 Qy 685 CGGCTTTCGCGGCTATGATTAAGAGCTGATTTGTCAGTTCCTACTCTCTAAAACTTT 744
 Db 751 CGGCTTTCGCGCACCAGACTGCGCGAGGTGCTGATCGCGGCTCTCTGCTCGAAGAACTTC 810
 Qy 745 GCGCTGTACAACGAGCGTGTGCGGCTTGTACTCTGTTGCTGCCGACAGTGAACCGTT 804
 Db 811 GGCATCTACCGGAGCGAAACGGGCATCTCGA---TCGCCATCGCGAGGGCGGGCGCG 867
 Qy 805 GATCGCGCATTCAGCCAAATGAAAGCGGCGATTTCGCGCTAACTACTCTAACCCACAGCA 864
 Db 868 GGCACGGTGCAGGCCAACCTCACTTCCTGAAACCGCGCAGAACTACTCTTCCCGCGGAC 927
 Qy 865 CACGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACGATGCGTTAGTGGGATTTGGGAA 924
 Db 928 CATGGCGCGGCTCGTGACCATGATCTCTCGAGGACGAGACGCTGAGCGCGGCTGGAAG 987
 Qy 925 CAAGAGCTGACTGATATGCGCAGCGTATTCAGCGTATGCGTCAAGTTGTTGCTCAATACG 984
 Db 988 GCGGAACTCGAGGAGGTGCGGCTCAACATGCTGACCGTGGCGCGCGCAGCTTCGCGATCG 1047
 Qy 985 CTGCGAGAAAAAGGCGCAAAACCGGACTTCAGCTTTATCATCAAAACAGAACGCGATGTT 1044
 Db 1048 CTGAGGCCGAGACCGGCTCGAACCGCTTCGGCTTCGTTGCGCGCAGCATTCGCGCATGTT 1107
 Qy 1045 TCCTTCAGTGGCTGACAAAGAAACAAAGTGTGCTGCTGGCGAAGAGTTTGGGTTATAT 1104
 Db 1108 TCGCGCTCGGATCACGCCCGCGAGGTGGAGCGGCTGCGGACCGGAGCAGCGGGTCTAC 1167
 Qy 1105 GCGGTTGCTTCTGTTGCGGTAATGTTGGCCGGGATGACACAGATACATGCGTCCGCTG 1164
 Db 1168 ATGGTGGCGATTCGCGGCTGAAACATCGCGGGCTGAAACCGGAGCGACCGTGCCTG 1227
 Qy 1165 TGCAGAGCGATTGTGGCAGTGTCTG 1188
 Db 1228 GCGCGCGGTGCGCAAGGTGCTG 1251

RESULT 8

```

US-10-979-821-3
; Sequence 3, Application US/10979821
; Publication NO. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 3
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Rhodobacter sphaeroides
US-10-979-821-3

Query Match          20.2%; Score 240; DB 8; Length 1260;
Best Local Similarity 51.7%; Pred. No. 1e-67;
Matches 602; Conservative 0; Mismatches 550; Indels 12; Gaps 2

Qy      Db      Qy      Db      Qy      Db      Qy      Db      Qy      Db      Qy      Db      Qy      Db      Qy      Db      Qy      Db      Qy      Db
25  CCTGCCGACCGAGTTCTGGGGCTTCTGGCGCATCTGTTCGTGCCGATGAACGTCCTCCGCGCAA 84
   |||||
100 CCCGGGACAGATCCTGCAACTGCATCCAGATGTTCCGGGAGATCGCGCGGGACAG 159
   |||||
85  ATTAACCTCGGGATTGGTGCTATAAAGATGAGACGGGCAAAACCCTGGTATGTACCTCGGCATT 144
   |||||
160 ATCGATCTGGGCGTGGGGCGTCTACAAGAACCCGACCGGGCTCACCCCGGTCATCGGGGCC 219
   |||||
145 GTGAAAAGGCTGACAGTAGTATCTGCTCGAAAAATGAAACCAACCAAAATTTACCTCGGCATT 204
   |||||
220 GTGAAGCGGCGGAGAAGCGGCTCTGGGAGGTGCAGACCAACCAAGACCTTACACCGGCGCTT 279
   |||||
205 GACGCGATCCCTGAATTTGGTTCGTGCACTCAGGAACGTGCTTTGGTAAAGGTAGCGCC 264
   |||||
280 GCCGACGAGCCGGCTACAAATGCCCGATGGCGAAGCTGATCCTCG-----CGGGC 330
   |||||
265 CTGATCAATGACAAAGTGTCTGCA CGGCAAGACTCTCGGGGGGGCACTCGGCGCATTCGCG 324
   |||||
331 CGGCTCCCGGCGGACCGGGGTGGCTCGGTCCCAACCCCGCGGCGCACGCGGCGGGTGGGT 390
   |||||
325 GTGGCTGCGCGATTTCTCGGCAAAAAATACCAAGGTTTAAAGGTGTGGGTGAGCAACCCA 384
   |||||
391 CAGGCGCTCGAGCTGATCCGATGGCTTCGCCGAGGCGCACCGCTCTGGATCTCGAACCGG 450
   |||||
385 AGCTGGCGGAACCAAGAGCGCTCTTTAACTCTGCAGGCTCTGGAAGTTCGTGAATACGCT 444
   |||||
451 ACCTGGCGGAACCATCTGTCGATCGTAATATCTCGGCATCCGATCGGGAATACCGC 510
   |||||
445 TATTATGATGGGAAAATCACTCTTTGACTTCGATGCATGATTAACAGCCTGAATGAA 504
   |||||
511 TATTTCGACGCGGAGACCGGGCGGCTCGATCGCGGAGGCATGATGAGAGATCTGCCCGAG 570
   |||||
505 GCTCAGGCTGGGACGTAGTGTCTTCCATGGGTGCTGCCATAACCAACCGGTATCCAG 564
   |||||
571 GTGAAGCGGCGGACGTGGTGTCTGCTGCA CGGCTGTCTGCCAACCAACCCGACCGGCGCAAC 630
   |||||
565 CCTAGCTGGAAACAATGGCAAAACACTGGCAAACTCTCTCGTTTGAGAAAAGGCTGGTTACCG 624
   |||||

```

Db 100 CCCGCGCAAGATCCTGCAACTGATCCAGATGTTCCGCGAGGATGCGCGCGGACAG 159
 Qy 85 ATTAACCTCGGGATGTGTCTATAAGATGAGACGGGCAAAACCCCGGTACTGACACAGC 144
 Db 160 ATCGATCTGGCGTGGCGGTCTACAAGACCCGACCGGGTCAACCCGGTCAATCGCGGCC 219
 Qy 145 GTGAAAGGCTGACAGTATCTGCTGAAATGAAACACCAACCAAAATTAACCTCGGCAAT 204
 Db 220 GTGAAGCGCGCGAAGACGGCTCTGGGAGGTGAGACCAACCAAGACTACACCGGCTT 279
 Qy 205 GACGGCATCTGAAATTTGGTGTGCTGCACTCAGGAATGCTGTTTGTAAAGGTAGCGC 264
 Db 280 GCCGACAGCGGGCTACAATGCCGCGATGGCGAGCTGATCTCTG-----CGGC 330
 Qy 265 CTGATCAATGACAAACGTGTGCGACCGGCAAGACTCCGGGGGCACTGCGCACTACGC 324
 Db 331 CGGCTCCCGCGACCGGTGGCTCGCTCGCCACCCCGCGGCAAGCGGTGCGT 390
 Qy 325 GTGGCTCCGATTTCTTGGGCAAAATAACAGCGTTAAGCGTGTGTGGGTAGCAACCA 384
 Db 391 CAGCGCTGCAAGCTGATCCGATGGCTCGCCCGAGGCCACCGTCTGGATCTCGAACCCG 450
 Qy 385 AGCTGGCGCAACCAAGAGCGTCTTTAACTCTGCAAGTCTGGAAGTTCGTGAATACGCT 444
 Db 451 ACCTGGCGCAACCAATCTGTGATCGTGAATATCTCGGATCCCGATGCGGGATACCG 510
 Qy 445 TATTATGATCGGAAATACACACTCTTGAATTCGATGCACTGATTAACAGCCTGAATGAA 504
 Db 511 TATTTCGACGCGAGACCGGCGCGTGCATGCGAGGGCATGATGAGGATCTGGCCAG 570
 Qy 505 GCTCAGGCTGGCGAGTGTGCTGTTCCATGCGTGTGCTGCCATACCCCAACCGGTATCGAC 564
 Db 571 GTGAAGCGCGGCGAGTGTGCTGCTGCAAGGCTGTGCTGCAACACCGCGCGCGCAAC 630
 Qy 565 CCTACGCTGGAACAATGGCAACACTGGCAACACTCTCCGTGGAAGGCTGGTTACCG 624
 Db 631 CCGAACCGGTGAGTGTGGCTGATGCGAGGCTGCGCGGACAGCGCGGTGCGG 690
 Qy 625 CTGTTGACTTCGTTACCAAGGTTTTGCGCGGTGCTGGAAGAGATGCTGAAGACTG 684
 Db 691 CTGATCAGCTCGCTTATCAGGCTTCGCGACGCGGTGCGAGATGATGCGCGCGACG 750
 Qy 685 CGCGCTTCGGGCTATGCAATAAGAGCTGATTTGCGAGTTCCTACTCTAAACATTT 744
 Db 751 CGGCTTCGGCCACCACTGCGCGAGTGTGATCGCGGCTCTCTCTCGAAGAACTTC 810
 Qy 745 GGCCTGTACACAGAGCGTGTGGGCTTGTACTCTGCTGCTGCGGACAGTGAACCGTT 804
 Db 811 GGCATCTACCGGAGCGCAAGGATCTCTGA--TCGCCATCGCGAGGCGCGGCGG 867
 Qy 805 GATCGCGCATTCAGCCAAATGAAGCGGCGATTCGCGCTAACTACTTAACCCACAGCA 864
 Db 868 GGCACGTTGCGGCGCAACCTCACTTCTGAACCGGAGAACTACTCTTCCCGCGGAC 927
 Qy 865 CAGGCGCTTCTGTTGTTGCAACATCTGAGCAACGATGGTTACGTGCGATTTGGGAA 924
 Db 928 CATGCGCGCGGCTCGTGACCATGATCTCGAGGACGAGACGCTGAGCGCGCATGGAAG 987
 Qy 925 CAAGAGCTGATGATATGCGCGACGCTATTACAGGATGCTGATGTTGCTCAATACG 984
 Db 988 CGGNACTCGAGGAGTGGGCTCAACATGCTGACATGCGCGCGGAGCTTGGCGATGCG 1047
 Qy 985 CTGAGGAAAGCGCAACCGGACTTCAGCTTTATCATCAACAGAAACGGGATGTTTC 1044
 Db 1048 CTGAGGCGGAGACCGGCTCGAACCGCTTCGGCTTCGTGGCGGAGCATCGCGCATGTC 1107
 Qy 1045 TCCTTCACTGCGCTGACAAAGAAACAAGTGTGCTGTGCGGCAAGAGTTTGGCGTATAT 1104
 Db 1108 TCGCGCTTCGGGATCAGCGCGCGGAGTGGAGCGGCTGGGACCGGAGCGGGGTCTAC 1167
 Qy 1105 CGGCTTCTGCTGCGTAAATGTGCGCGGAGTACACAGATAACATGGCTCGCGCTG 1164

Db 1168 ATGGTGGCGATTCGCGGCTGAACATCGCGGGCTGAACCGGACCGGTGCGGTGCTG 1227
 Qy 1165 TCGGAAGCGATTTGTGCACTGCTG 1188
 Db 1228 GCGCGCGCGTGGCCAAAGGTGCTG 1251
 RESULT 10
 US-11-136-527-2278
 ; Sequence 2278, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031996-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2278
 ; LENGTH: 2343
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; US-11-136-527-2278
 Query Match 18.6%; Score 221.6; DB 12; Length 2343;
 Best Local Similarity 50.1%; Pred. No. 1.5e-61;
 Matches 580; Conservative 9; Mismatches 563; Indels 6; Gaps 2;
 Qy 25 CTGCGCGACCGGATTTCTGGCGCTGGCGATCTGTTTCGTGCCGATGAACGTCGCCGCAAA 84
 Db 207 CTTCCAGATCCCATCTCTGGAGTGACCGAGCTTCAGWAGAGATACCAACAGCAMGAAG 266
 Qy 85 ATTAACCTCGGATTTGTGTCTATAAGATGACACGGGCAAAACCCCGGTACTGACGACG 144
 Db 267 ATGAACCTGGAGTTGGTGTCTACCGGACGATTAACGGAAGCTTACGTGCTCCCGAGT 326
 Qy 145 GTGAAAGGCTGAAACAGATATCTGCTCGAAATGAAACCAACCAAAATTAACCTCGGCAAT 204
 Db 327 GTTCGGAAGGACGAGCCCGAGATTTCTGGGAAATTTTGGWCAAGAAATACCTACCCATC 386
 Qy 205 GACGGCATCTGTAATTTGTGCTGCACTGAGAACTGCTGTTTGTAAAGGTAGGCGCC 264
 Db 387 GGGGGAATGCTGATTTTGTAAAGGTTCTGCAAGACTGGCCCTGGGCGAGAACAGCGAA 446
 Qy 265 CTGATCAATGACAAACGCTGCTCGCAGGACAGACTCCGGGGGCACTGGCGCACTACGC 324
 Db 447 GTGTTGAAAGCGCGCGGTTTGTAACTGTGACAGCCATTTCCGGGACTGGAGCCTTGAGG 506
 Qy 325 GTGGCTCGCGATTTCTTGGCAAAATAACAGCGTTAAAGCG---TGTGTGGGTGAGCAAC 381
 Db 507 GTGCGGACCGAGCTTTCTGCAAGATTTTAAAGTTCAGCCGAGATGCTCTTCTGCCCAAA 566
 Qy 382 CCAAGCTGGCGCAACCATTAAGAGCGCTTTAACTCTGCAAGTCTGGAAGTTCGTAATAC 441
 Db 567 CCATCTCTGGGAAACCAACAGCCCATCTTCAGGGATGCCGCGATGCAAGGTTAT 626
 Qy 442 GCTTATTATGATCGGAAATACACTCTTGAATTCGATGCACTGATTAAACAGCCTGAAT 501
 Db 627 CGATCTATGACCCCAAGACTTGGCGTTTGAATCTCTGAGGCTTAGAAGACATATCA 686
 Qy 502 GAAGCTCAGGCTGGCGACGTAGTGTTCATGGGTGTCCTAATACCAACCGGTATC 561
 Db 687 AAATCCGAGCAGAGTGTCTCTCTTGTGACGCTGCGCTCACACCCACCGGCGTG 746
 Qy 562 GACCTTACGCTGAAACATAGGCAACACTGGCAACACTGCGCAACTCTCCGTTGGAAGAGCTGTTA 621
 Db 747 GACCCGCTCCAGAGCAGTGAAGAGAAATGGCGCGGTGGTGAAGAAAAAATCTCTTC 806
 Qy 622 CGGCTGTTGACTTCGCTTACCAGGGTTTTCGCCCGTGG---TCTGGAAAGAGATGCTGAA 678

```

Db 807 GCATTTTGAATGCGCTTACCAAGGCTTTGCGAGCGGATGGTGAAGAGCGCTGG 866
Qy 679 GGACTGCGCGCTTTGCGGCTATGCTAAGAGCTGATTTGTCAGGTTCTTACTCTAA 738
Db 867 GCCGTGCGGCACTTCAATGAGCAGGATCAATGCTGCTCTGCAATCTTATGCCAAG 926
Qy 739 AACTTTGGCCTGTACAACGAGCGTGTGGCGCTTGTACTCTGTTGCTGCGCAGAGTAA 798
Db 927 AACATGGCGCTGTACGGTGAGCGTGTGGAGCCTTCACTGTGTGTGCAAAAGATGCA 986
Qy 799 ACCGTGTATGCGGCACTTCAAGCAATGAAGCGGCGATTCGGGCTAATCTTAAACCA 858
Db 987 GAAGCAAAAGGGTGGAGTCAAGCTGAAGATCTCTGATCCGCGCTTGTATTTCCAAAC 1046
Qy 859 CCAGCACACGCGCTTCTGTTGTTGTCACCATCTCTGAGCAAGATGCGGTAGCTGCGATT 918
Db 1047 CCTCTCAATGAGCCCGGATCGCCCAACCATCTCTGATCTSSAGACKSCGGAAGCAA 1106
Qy 919 TGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACAGCGTATGCGTCAGTTGTCG 978
Db 1107 TGGTTGAGGAGGTGAAGGATGCTGACCGCATCATCAGCATGAGGACCGAGTTGTC 1166
Qy 979 AATAGCTGCGAGGAAAAAGGCGCAAAACCGGACTTTCAGCTTATCATCAAAACAGAAC 1038
Db 1167 TCCAACTGAAGAAAGAGGCTCGTCCCAACTGGCAGCACATCAGCGACAGATCGC 1226
Qy 1039 ATGTTCTCTTCACTGCGCTGACAAAGAAACAGTGTGCTGCGGAGAGTTGGC 1098
Db 1227 ATGTTCTCTTCACTGCGCTGAAAGCTTGAAGCTGAGCAGTGGAGCGGCTGACCAAG 1286
Qy 1099 GTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCGGAGTGAACACAGATAAATGCT 1158
Db 1287 GKCTACATGACAAAGATGTTGCTGCTCTGKGGCGGGTCACTCTGGCAATGTGGC 1346
Qy 1159 CGCTGTGCGAAGCGATT 1176
Db 1347 TAMCTGGCCACGCCATT 1364

```

RESULT 11

```

US-10-979-821-1
; Sequence 1, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 1
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Sinorhizobium meliloti
US-10-979-821-1

```

Query Match 17.3%; Score 206.2; DB 8; Length 1170;
 Best Local Similarity 50.8%; Pred. No. 1.1e-56;
 Matches 580; Conservative 0; Mismatches 543; Indels 18; Gaps 3;

```

Qy 1 ATGTTTGAGAACATTACCGCGCTCCTGCGCAGCCGATTTCTGGGCTCGCGGATCTGTTT 60
Db 1 ATGTTTGAGCGCTCGCGCCGCAAGCGACGATCCCTTGTCTTCTCTGATCGGCTGTC 60
Qy 61 CGTCCGAGATGAAGTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGAG 120
Db 61 AGGAAGATGAGCGCGCGGAAAGGTTCGATCTCGCGGTAGAGTCTATCGGACGAGACC 120
Qy 121 GGCAAAACCCCGGTACTGACAGCGTGAAGAGCTGAACAGTATCTCTCGTGAAGTAA 180
Db 121 GAGCGACGCCGATCTTCGGGCGCTCAAGCGCGGAAAGCGGCTTCTCGAAACAG 180
Qy 181 ACCACCAAAATTAACCTCGGCAITTCAGCGCATTCCTGAAATTTGGTGTCTGCACTCAG 240
Db 181 GACAGCAAGGCTATATCGGCGCGAGGGGACCTCGTCTTCTCGATCGGCTCTGGGAA 240
Qy 241 CTGCTGTTTGTAAAGTAGCGCCTCATTAATGAACAAACGTGTGCGACGCGACAGACT 300
Db 241 CTCGT-----CGGCGCGACACGATCGAGCGGAGCCATGTTGCGGGGCTCCAGAG 291
Qy 301 CCGGGGGCACTGGGCACTAGCGGTGGCTGCGGATTTCTTGGGCAAAATTAACAGCGTT 360
Db 292 CCGGGCGGCTCCGGCGGCTCCGTTTGGCGCGGACCTCATCGC---CCGATGGCGGCG 348
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAAGAGCGTCTTTAACTCTGCA 420
Db 349 CGAGCGATCTGCTCGGCTCGGAGTGGCGGCAACCAAGCGCGGATCTTCAAGCGGCG 408
Qy 421 GGTCTGGAAGTTCGTGAATAGCTTATTAATGATCGGAAATCAACACTTTGACTTCGAT 480
Db 409 GGGCTCGATATCGCACCTACGACTTCTTCGACATTCGTCGCGAGTCGTCATCTTCAT 468
Qy 481 GCATGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCAATGGCTGC 540
Db 469 AATCTGGTGAAGCGCTGGAAGCGCGCCGATCCGGCGATGCGGTGCTGCTCATGCAAG 528
Qy 541 TGCCATAACCAACCGGTATCGACCTTACGCTGGAACAAATGGCAAAACATCGGCACAACT 600
Db 529 TGCCACAACCGGACCGGGGCTCTGAGCGGAACACATGATGAGATCGCGCGCTG 588
Qy 601 TCGGTTGAGAAAGCTGTTTACCGCTGTTGACTTTCGCTTACACAGGGTTTTCGCGTGT 660
Db 589 GTGCGCGAGCGCGCTGCTGCGCTCGTCGATCTCGCTATCAGGGGTTTCGCGCGGCG 648
Qy 661 CTGGAAGAAGTGTGAAGGATCGCGCTTTCGCGGCTATGCAATAAAGAGCTGATGTT 720
Db 649 CTCGACCAAGATGTGCGGGGCTCCGCACTTCTCGCGGTGCTCCGGAAGCGCTCGTC 708
Qy 721 GCCAGTTTCTTCTAAATACTTTGGGCTGTACAAACGAGCGTGTGGGCTTGTACTCTG 780
Db 709 CGGGTTTCTGCTCGAAAGTCTTCGGGCTTATCGGAGCGCGGCGGCGGATCTTCGCG 768
Qy 781 GTTGTCTCGGACAGTGAACCGTTGATCGCGCATTCAGGCCAAATGAAGCGCGGATTCGC 840
Db 769 CGGACCACTCGACTGCTCGGCGGACAGGGTGGCTCAAACTCGCGGGCTCGCACGC 828
Qy 841 GCTAACTACTTAACCCACAGCACAGGGGCTTCTGTTGTTGGCCACCATCTGAGCAAC 900
Db 829 ACCAGCTATTTCATGCGCGCGGATCAGCGGCGAGCGCTGTCGCGGACGATCTCTTAC 888
Qy 901 GATGCGTTAGTGGATTTGGGAACAGAGTGAATGATGCGCGGAGCTTATTCAGCGT 960
Db 889 CCGGAACCTCAGGCGCGACTGAGCGAGGAGCTCGAGAGCTCGGGCTCAGGATGACGGCG 948
Qy 961 ATGCTCAGTTGTTGCTCAATACGCTGCAAGAAAAAGCGCAAAACCGGACTTTCAGCTTT 1020
Db 949 CTCGGCGGTGCTTGGCGAGGAGCTCGGACCGCTGGCAGAGC-----TCGGCGCA 1002
Qy 1021 ATCATCAAAACAGAACGGCATGTTCTCTCAGTGGCTCTGACAAAGAACAAAGTGTGCGT 1080
Db 1003 GTCCCGATCAGGAGGCGATGTTCTCCATGTGCGGCTTTCGGAAGCGGAGGTTATGCGG 1062

```


Qy 1081 CTGCGCGAAGAGTTTGGGTATATGTCGTTGCTTCTGTCGCGTAAATGTGGCGGGATG 1140
Db 1063 CTCAGGACCGAGCAGCGGATCTATATGCGGCGATCCGCGCGCATCAACATCGCGGGCTG 1122

Qy 1141 A 1141
Db 1123 A 1123

RESULT 12
US-11-114-922-1
; Sequence 1, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 1
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Sinorhizobium meliloti

Query Match 17.3%; Score 206.2; DB 12; Length 1170;
Best Local Similarity 50.8%; Pred. No. 1.1e-56;
Matches 580; Conservative 0; Mismatches 543; Indels 18; Gaps 3;

Qy 1 1 ATGTTTGAACATACCGCGCTCTGCGGACCGGATTCCTGGGCGCTGGCGGATCTGTTT 60
Db 1 1 ATGTTGACGCGCTTCCGCGCGCAAGCGGATCCCTTGTCTTTTCTGATCGGCGCTGTT 60

Qy 61 CGTGCCGATGAAGTCCCGGCAAAATTAACCTCGGGATTTGTTCTATAAGATGAGACG 120
Db 61 AGAAGGATGAGCCCGCGGAAAGTGTCTCGGCGTAGAGTCTATCGGACGAGCC 120

Qy 121 GCGAAAAACCCGGTACTGACGAGCGTGAAAAAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
Db 121 GGACGACGCGGATCTTCCGGCGGTCAAGGCGCGGAAAAAGCGGCTTCTCGAAAAACAG 180

Qy 181 ACACCAAAATTAACCTCGGATTCAGCGGATTCCTGAAATTTGTCGCTGCACTCAGGAA 240
Db 181 GACAGCAAGGCTATATCGGCGCGCAAGGAGGACCTGCTCTTTCTCGATCGGCTCTGGGAA 240

Qy 241 CTGCTGTTTGTAAAGTGTAGCGCTCTGATCAATGACAAAGCTGCTCGCAGCGGACAGCT 300
Db 241 CTCGT-----CGGCGGCAACAGATGAGCGAGCCATGTTGGGGGCTCCAGACG 291

Qy 301 CCGGGGGGCACTGGCGCACTACGGGTGGCTGCGGATTTCTTGGCAAAAAATACACGGGTT 360
Db 292 CCGGGGCTCCGGCGGCTCGGTTTGGCGCGGACCTCATCGC---CCGATGGCGGC 348

Qy 361 AAGCGTGTGGGTGAGCAACCCAGCTGGCGGCAACCAATAGACGCTCTTTAACTCTGCA 420
Db 349 CGAGGCATCTGGCTCGGCTCGCGGCTGGCGGCAACCAACGCGCGGATCTTCAAGCGGCC 408

Qy 421 GGTCTGGAAGTTCGTGAATGCTTATATGATCGGGAATACACTCTGACTTCGAT 480
Db 409 GGGCTCGATATCGGCACCTAGACTTCTTCGACATTCGTCGACGTGGTCACTTCGAT 468

Qy 481 GCACGTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTGCTTCCATGGCTGC 540
Db 469 AATCTGGTGGCGGCTGGAGGCGCGCGCATCCGCGCATCGGTGCTGCTGCTCATGCAAGC 528

Qy 541 TGCCATAACCCAACCGGTATCGACCCCTACGCTGGAACAAATGGCAACACACTGGCACAACTC 600
Db 529 TGCACAAACCCGACCGCGCGGTCTCTGAGGAAGCACATGGATGGAGATCGCGCGCTG 588

Qy 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACCGGTTTACAGGGTTTGGCCG 660
Db 589 GTGGCGGAGCGCGGCTGCTGCGGCTCGTCGATCTCGCTTATCAGGGGTTTCGCGCGCGC 648

Qy 661 CTGGAAGAAGATCTGAAGGACTGCGGCTTTTCGCGGCTATGATGAAGAGCTGATGTT 720
Db 649 CTCGACCAAGATGTGCGGCGCTTCGCGCATCTTCTCGGCGTGGTCCCGGAAGCGCTGCTC 708

Qy 721 GCCAGTTCTACTCTAAAAAATTTGGCTGTCAACAGAGCGTCTTGGCGCTTGTACTCTG 780
Db 709 GCGGTTTCTCTCGAAGTCTTTCGCGCTTATCGCGAGCGCGCGGCGGCGATCTTCGCG 768

Qy 781 GTTGTGCGGACAGTGAACCGTTGATCGGCAATTCAGCCAAATGAAGGGCGGATTCGC 840
Db 769 CGGACCAAGCTCGACTGCTCGCGGACAGGGTGGCTCAAACTCGCGGCGCTCGCACGC 828

Qy 841 GCTAACTACTCTAACCCACGACGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
Db 829 ACCAGCTATTCCATGCGCGCGGATCACGGCGAGCGCTGTCGCGACGATCCTTGAGCAG 888

Qy 901 GATCGTTAGTGCAGTTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
Db 889 CCGAACTAGGCGCGACTGGACGGAGGAGCTCGACAGTATGCGGCTCAGGATGACGGGC 948

Qy 961 ATCGCTCAGTTGTTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCGACTTCAGGTTT 1020
Db 949 CTCGCGGCTGCTTCCGAGGAGCTCCGACCCGCTGGCAGAGCC-----TCGGCGCA 1002

Qy 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGCAAAAGAACAACTGCTGCGT 1080
Db 1003 GTCCCGCATCAGAGGCGATGTTCTCCATGCTGCCGCTTTCGAAGCGGAGGTATCGG 1062

Qy 1081 CTGCGGAAAGTTCGCGTATATGCGGTTGCTTCTGTCGCTAAATGTGGCGGGATG 1140
Db 1063 CTGAGACGAGCAGCGCATCTATATGCGGCGATCGCGCGCATCAACATCGCGGGCTG 1122

Qy 1141 A 1141
Db 1123 A 1123

RESULT 13
US-10-979-821-7
; Sequence 7, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Leishmania major

US-10-979-821-7

```

Query Match      12.2%; Score 145.8; DB 8; Length 1239;
Best Local Similarity 48.6%; Pred. No. 6.8e-37;
Matches 492; Conservative 0; Mismatches 512; Indels 9; Gaps 3;

Qy 64 GCCGATGACGTCGCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACGGGC 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 94 GCCGCTGCAAGGGCCCCAAGGCCAACCTCGTCATTGGTGCTACCGGACGAGCAGGGC 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 AABACCCCGGTACTACACAGCGTGAAAGAGCTGAAGAGCTGAAGTATCTGCTGAAATGAAC 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 154 CGTCCCTATCCGCTACGCGTGGTCGCAAGCTGAGCAGCTCTCTTGGACATGAATCTC 213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 184 ACCAAATTAACCTCGGCAATGACGCGATCCCTGAATTTGGTCGCTGACCTCAGAACTG 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 214 GACTACGATGATACCTACCTCTCGGCTACAGCCCTTCATCGATGAGCGGTAAAGATT 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 CTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAGCTGCTCGCAGCGCACAGCTCCG 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 274 ATCTACGGCAA--TACCGTCAAGCTGAGGAACTCTGCTCGGCTGAGGCTGAGCGGC 330
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 304 GGGGCACTGCGGCACTACGCGTGGTCCGATTTCTGCGCAAAATACAGGCTTAAG 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 331 ACCGGTGTCTCTCTCTCGGGCGAAGCTGTGACTCGCGTCTTCGACGCTGAGACGAGC 390
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 364 CGTGTGGGTGAGCAACCCCAAGCTGCGCAATGAGAGCGTCTTTAACTCTGCGAG-- 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 391 CCCATCTACCTTTCCGACCCCAAGCTGCGGCTACCGGCTGAGGCTGAGGCTGCTGCG 450
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 422 -GTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAATACACACTCTTTGACTTCGAT 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 451 TGGAGAAACATCTGACGTCAGCTACTACGACCCCAAGCGGTCAAGCTGAAATTCGAG 510
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 481 GCATGATTAACAGCTGAGCTGATCAATGACAAAGCTGCTCGCAGCGCACAGCTCCG 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 511 GGCATGAAGAAAGACATTTCTGCGCGCGCGGACGCGCTCCGTTGTTCAATTCGACACAGTGC 570
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 541 TGCCATTAACCAACCGGTATCGACCTACGCTGGAACAAATGCGCAAAACACTGCGCAACTC 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 571 GCGCACAAACCCCAACCGGCTGGAACCGCTGCGAGGACGATGGAACGAGATCGCGTCACTG 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 601 TCCGTTTGAGAAAGGTGTTTACCCTGTTTACCTTTCGCTTACCAGGGTTTTGCG--CCGT 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 631 ATGCTGCCAAGACCATCAGGTGTTCTTTCGCTCCGCTACCAAGGCTATGCGAGCGGC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 658 GGTCTGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGGCTATGCAATAAAGAGCTGATT 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 691 AGCCTCGACACGAGCGGCTATGCTGCGCGCTGTTTGGCCCGCGGCACTCGAGGTACTG 750
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 718 GTTCCAGTTCTTACTCTTAAACCTTTGGCCCTGTACACGAGCGTGTGGCGCTTGACT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 751 CTGGCGCAGTGTCTCTCAAGAAATGCGCTTGTACAGCGAGCGTGGAGGACGCTGTCG 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 778 CTGGTTGTCGCGCAGTGAAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATT 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 811 CTGCTCTCAAGACACAGCAAGCGCGGATGTAAGAGCTGTGATGATTCGCTGATC 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 838 CGCGCTAATCTACTTAACCCACACAGCAACCGCGCTTCTGTTGTTGCCAATCCTGAGC 897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 871 CGTGAGGAGTACAGTGGCGCCCAAGCCCAAGCTGCGCGCTTAGCCCACTTAATCCTGAGC 930
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 898 AACCATGGTTACGTGCGATTGGGACACAGAGCTGATGATGCGCCAGCGTATTTCAG 957
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 931 AACCAACGAACCTGCGAAAGAGTGGAGGACAGAGTATCAGCCATGGCAGAGCGCATCCGT 990
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 958 CGTATGCGTCAAGTTGTTGCTCAATACGCTGAGGAAAGCGCAAAACCGGAGCTTCAGC 1017
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 991 ACGATGCGCGCACCGGTGACGACGAGCTGTGCGCTGCGAGACGCCGCGGAGCTGGAA 1050
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1018 TTTATCATCAACAGAGACGGCATGTTCTCTTCAGTGGCCTTGACAAAGAAC 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1051 CATGTCAATTAACAGATTGGCATGTTTCTCTCTCGGGCTGTCAAAGGGCA 1103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
US-11-114-922-7
; Sequence 7, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Leishmania major
; US-11-114-922-7

Query Match      12.2%; Score 145.8; DB 12; Length 1239;
Best Local Similarity 48.6%; Pred. No. 6.8e-37;
Matches 492; Conservative 0; Mismatches 512; Indels 9; Gaps 3;

Qy 64 GCCGATGACGTCGCGGCAAAATTAACCTCGGGATTGGTGTCTATAAAGATGAGACGGGC 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 AABACCCCGGTACTGACGCGTGAAAGAGCTGAAGAGCTGAAGTATCTGCTGAAATGAAC 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 154 CGTCCCTATCCGCTACGCGTGGTCCGCAAGCTGTGACTCGCGTCTTCGACGCTGAGACGAGT 213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 184 ACCAAATTAACCTCGGCAATGACGCGATCCCTGAATTTGGTCGCTGACCTCAGAACTG 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 214 GACTACGATGATACCTACCTCTCGGCTACCGGCTTACCGGCTTTCATCGATGAGCGGTAAAGATT 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 CTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAAGCTGCTCGCAGCGCACAGACTCCG 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 274 ATCTACGGCAA--TACCGTCAAGCTGAGGAACTCTGCTCGGCTGAGGCTGAGCGGC 330
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 304 GGGGCACTGCGGCACTACGCGTGGTCCGATTTCTGCGCAAAATACAGGCTTAAG 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 331 ACCGGTGTCTCTCTCTCGGGCGAAGCTGTGACTCGCGTCTTCGACGCTGAGACGAGC 390
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 364 CGTGTGGGTGAGCAACCCCAAGCTGCGCAATGAGAGCGTCTTTAACTCTGCGAG-- 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 391 CCCATCTACCTTTCCGACCCCAAGCTGCGGCTACCGGCTGAGGCTGAGGCTGCTGCG 450
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 422 -GTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAATACACACTCTTTGACTTCGAT 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 451 TGGAGAAACATCTGACGTCAGCTACTACGACCCCAAGCGGTCAAGCTGAAATTCGAG 510
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 481 GCATGATTAACAGCTGAGCTGATCAATGACAAAGCTGCTCGCAGCGCACAGCTCCG 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 511 GGCATGAAGAAAGACATTTCTGCGCGCGCGGACGCGCTCCGTTGTTCAATTCGACACAGTGC 570
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 541 TGCCATTAACCAACCGGTATCGACCTACGCTGGAACAAATGCGCAAAACACTGCGCAACTC 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 571 GCGCACAAACCCCAACCGGCTGGAACCGCTGCGAGGACGATGGAACGAGATCGCGTCACTG 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 601 TCCGTTTGAGAAAGGTGTTTACCCTGTTTACCTTTCGCTTACCAGGGTTTTGCG--CCGT 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 631 ATGCTGCCAAGACCATCAGGTGTTCTTTCGCTCCGCTACCAAGGCTATGCGAGCGGC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 658 GGTCTGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGGCTATGCAATAAAGAGCTGATT 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 691 AGCCTCGACACGAGCGGCTATGCTGCGCGCTGTTTGGCCCGCGGCACTCGAGGTACTG 750
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 718 GTTCCAGTTCTTACTCTTAAACCTTTGGCCCTGTACACGAGCGTGTGGCGCTTGACT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 751 CTGGCGCAGTGTCTCTCAAGAAATGCGCTTGTACAGCGAGCGTGGAGGACGCTGTCG 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 778 CTGGTTGTCGCGCAGTGAAACCGTTGATCGCGCATTCAGCCAAATGAAAGCGCGGATT 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 811 CTGCTCTCAAGACACAGCAAGCGCGGATGTAAGAGCTGTGATGATTCGCTGATC 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 838 CGCGCTAATCTACTTAACCCACACAGCAACCGCGCTTCTGTTGTTGCCAATCCTGAGC 897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 871 CGTGAGGAGTACAGTGGCGCCCAAGCCCAAGCTGCGCGCTTAGCCCACTTAATCCTGAGC 930
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 898 AACCATGGTTACGTGCGATTGGGACACAGAGCTGATGATGCGCCAGCGTATTTCAG 957
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 931 AACCAACGAACCTGCGAAAGAGTGGAGGACAGAGTATCAGCCATGGCAGAGCGCATCCGT 990
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 958 CGTATGCGTCAAGTTGTTGCTCAATACGCTGAGGAAAGCGCAAAACCGGAGCTTCAGC 1017
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 991 ACGATGCGCGCACCGGTGACGACGAGCTGTGCGCTGCGAGACGCCGCGGAGCTGGAA 1050
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1018 TTTATCATCAACAGAGACGGCATGTTCTCTTCAGTGGCCTTGACAAAGAAC 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1051 CATGTCAATTAACAGATTGGCATGTTTCTCTCTCGGGCTGTCAAAGGGCA 1103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Qy 718 GTTCCAGTTCTTAACTTTGGCTGTACACGAGCGTGTGGCGTTGACT 777
Db 751 CTGGCGAGTGTCTCAAGAACATGGGCTTTGACAGAGCGTGCAGGACGCTGTGC 810
Qy 778 CTGGTTGTCGCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATT 837
Db 811 CTGCTCTCAGGACACAGACGCGCGGATGTAAGAGCGGTGATGCTGATC 870
Qy 838 CGCGCTAACTACTTAACCCACACAGCACAGCGCGTCTCTGTTGTGGCCACCATCCTGAGC 897
Db 871 CGTGAGGATACAGTGTCCCGCCACAGCCAGCGTCCCGCTTAGCCACCTAAATCCTGAGC 930
Qy 898 AACCATCGTTCAGTGTGGATTGGGAACAAGAGTGAATGATGCCAGCGGTATTCAG 957
Db 931 AACAACTGCGGAAAGGAGTGGGAGGAGGATATCAGCCATGGCAGAGCGCATCCGT 990
Qy 958 CGTATGCGTCAAGTTGTTTCGTAATACGCTGAGGAAAAAGCGCAAAACCGCGACTTCAGC 1017
Db 991 ACGATGCGCGCACCGTGTACGACGAGTGTGCGCCTGCAGAGCGCCGGGAGCTGGGAA 1050
Qy 1018 TTTATCATCAACAGACGCGCATGTTCTCTTTCAGTGGCTGTGACAAAGAAC 1070
Db 1051 CATGTCATTAAACAGATTGGCATGTTTCTTCTTCTCGGGTGTCAAAGGGCA 1103

RESULT 15
US-10-932-182A-82350
; Sequence 82350, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82350
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82350
```

```
Query Match 12.2%; Score 145.2; DB 7; Length 1251;
Best Local Similarity 48.4%; Pred. No. 1.1e-36;
Matches 572; Conservative 0; Mismatches 583; Indels 27; Gaps 5;

Qy 2 TGTTTGAGAACATTACCGCCCTCTCTGCCGACCGGATTTGGGCGCTGGCGGATCTGTTTC 61
Db 56 TGTTCAATAACATCGAATTGTGCGCCCTGTATGTCCTTTTGGTATTAAAGCAAGGTACG 115
Qy 62 GTGCCGATGAACGTCCCGGCAAAATTAACCTCGGGATTTGGTCTTAAGATGAGACGG 121
Db 116 GGCAAGATCAACGTCTACCAAGGTGACCTTTGGGTATCGGGGCTTACAGAGACGACAACG 175
Qy 122 GCAAAACCCGGTACTGACCAAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 181
Db 176 GTAAACCATGGGTCTTGCCAAAGTGTAAAGCCGCGAAAGCTTAATTCATACACAGCT 235
Qy 182 CCACGAAAAAT---TACCTCGGCATTGACGGCATCCCTGAAATTTGGTCTGCACTCAGG 238
Db 236 CCTACACCATGAATACCTCGGTATTACCGGTCTGCCAAGTTTGACATCTAACCGCCCA 295
Qy 239 AACTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCACGGCACA 298
Db 296 AGATCATCTTTCGGTACGCAATCCGATGCTTGCAGGAAGACAGAGTAAATCTCAGTACAAT 355
Qy 299 CTCGGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACCAGC 358
```

```
Db 356 CACTGTCTGTGTAACGGGTGCTCTTCATATATTTGCGAAGATTTTTTCAAAATCTCTCCAG 415
Qy 359 TTAAGCGTGTGGGTGAGCAACCAAGCTGGCGCAACCATAGAGCGTCTTTAACTCTG 418
Db 416 ATAACCTGGTCTATTTGTCTAAGCCTACTTGGGCCAACCAACATGGGCATTTTGAAGATC 475
Qy 419 CAGGTCTGGAAGTTCTGTAATACGCTTATATGATCGGAAAAATCACACTCTTTGACTTCG 478
Db 476 AAGCTTTGAACAGCGGACCTTACCTTACTGGGCCAACGAACTAAGTCTTTTGGACCTAA 535
Qy 479 ATGCACTGATTAACAGCCTGAATGAAGCTCAGCTCGCGACGTAGTGTCTTCCATGGCT 538
Db 536 ACGGCTTCTTAATGCTTATTTCAAAAAGCTCCAGAGGCTCCATTTTCGTTCTGCACTCTT 595
Qy 539 GCTGCCATTAACCAACCGGTATCGACCTACGCTGGAACAACTAGTCTTTTGGACCAAC 598
Db 596 GCGCCCATTAACCAACCTGGTCTGGACCTACTAGTGAACAACTGGGTTCAAACTGTTGATG 655
Qy 599 TCTCCGTTGAGAAAGCTGGTTACCGCTGTTTACCTTTCGCTTACAGGGTTTTGGCCGTG 658
Db 656 CTATCGCTCAAAAACCAACATCGCTTATTTGACCGCTTACCAAGGTTTGGCACTG 715
Qy 659 G---TCTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCATPAAGAG--- 711
Db 716 GAGATTTGGACAAAGGATGCTATGCTGCTAGTGTGGAGAAAGCTTTTCGACGCTCT 775
Qy 712 -----CTGATTTGTCAGTTCCTACTCTTAAACCTTTGGCCCTGTACACAGCGGTGTTG 766
Db 776 CTCCTGCTCTTGTCTGTCTAGTCTTTTGGCAAGAACCGCGGTATGTACGGGTGAGCGGTAG 835
Qy 767 GCGCTGTACTCTGCTGTTGCT-----GCCACAGTGAACCGCTTGTATCGCGCATTTCA 817
Db 836 GTTGTTCATCTAGCACTTACAAAACAGCTCAAACAACTATAAGACCTTGTCTGTTA 895
Qy 818 G---CCAAATGAAGCGGCGATTCGCGCTAACTACTCTAAACCAACGACACACGGCGCTT 874
Db 896 CATCTCAATTTGGCCAAATCATTTGTTAGTGAAGTGTCCAACCCACCGGCTACGGGCTA 955
Qy 875 CTGTTGTCGACCATCTCTGAGCAACGATCGTTAGTGGGATTTGGGAACAAGAGCTGA 934
Db 956 AGATTGTCGCTAACTGTTTGGAAACGCCAGAAATTAACGGAACAGTGGCAACAGGATATGG 1015
Qy 935 CTGATATGCGCCAGCGTATTCAGCGTATGCTGCTGCTGCTCAATACGCTGCAGGAAA 994
Db 1016 TTACCATGCTCCAGAAATTACGAAATGAGGCAACGCAATTAAGAGACCATTTAGTCAAGT 1075
Qy 995 AAGCGCAAAACCCGCGACTTTCAGCTTTTATCATAAACAGAACCGCATGTTCTCTTCAAGT 1054
Db 1076 TGGGCACTCTGCGCACTGGGATCATATAGTAAATCAATGCGGGATGTTCTCTTTTACAG 1135
Qy 1055 GCCTGACAAAGAACAGTGTGCTGCTGCGCGCAAGAGTTTGGCGTATATGCGGTTGCTT 1114
Db 1136 GATTGATCTCTCAATGTTTAAACGACTTTGAAGAAAAACCCACGCGAGTTTACTTGGTTGCC 1195
Qy 1115 CTGCTGCGTAAATGTTGGCGGATGACACCATTAACATG 1156
Db 1196 CAGGTAGAGCTTCTATTTGCTGGATTGAATCAAGAAACGTTGG 1237
```

Search completed: March 15, 2006, 14:54:58
Job time : 1440 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 17:56:20 ; Search time 40 Seconds
(without alignments)
952.546 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITRAPADPILGLDLF.....VAGTMDNMAPLCEAIVAVL 396
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2045	100.0	396	1 XNECD	aspartate transami
2	2037	99.6	396	2 A85619	aspartate aminotra
3	2037	99.6	396	2 C90755	aspartate aminotra
4	1778	96.7	396	2 AD0616	aspartate aminotra
5	1775	86.8	396	2 AD0172	aspartate transami
6	1329	65.0	396	2 I64132	aspartate transami
7	1315	64.3	413	2 A82217	transaminase (EC 2
8	1311	64.1	397	2 B81915	aspartate transami
9	1310	64.1	397	2 C81188	aspartate transami
10	982	48.0	398	2 B83252	probable amino aci
11	910.5	44.5	399	2 D83535	aromatic-amino aci
12	909	44.4	400	2 C82855	aromatic-amino aci
13	909	44.4	407	2 JC5124	aspartate transami
14	908	44.4	409	2 S53303	aspartate transami
15	904	44.2	405	2 T14311	aspartate transami
16	891.5	43.6	394	2 T06136	aromatic-amino aci
17	885	43.3	456	2 T06136	aspartate transami
18	879	43.0	418	2 S46315	aspartate transami
19	874	42.7	453	2 T04646	aspartate transami
20	873.5	42.7	404	2 G82449	amino acid biosynt
21	871	42.6	449	2 T48511	aspartate transami
22	863	42.2	463	2 S33528	aspartate transami
23	861	42.1	454	1 XNVLB	aspartate transami
24	861	42.1	455	2 S46315	aspartate transami
25	861	42.1	457	2 S65675	aspartate transami
26	859.5	42.0	397	2 AD0040	aromatic-amino aci
27	843.5	41.2	397	1 XNECY	aromatic-amino aci
28	842.5	41.2	397	2 H86098	tyrosine aminotran
29	842.5	41.2	397	2 D91258	tyrosine aminotran

30	839.5	41.1	397	2 S71928	aromatic-amino-aci
31	838.5	41.0	397	2 AG1016	aromatic-amino-aci
32	838	41.0	397	2 H81054	aromatic-amino-aci
33	832	40.7	397	2 G81821	aromatic-amino-aci
34	832	40.7	453	2 S39928	aspartate transami
35	829	40.5	453	2 S39927	aspartate transami
36	829	40.5	463	2 S39925	aspartate transami
37	827	40.4	463	2 S39926	aspartate transami
38	818	40.0	389	2 B47094	aromatic-amino-aci
39	812	39.7	430	1 XNPGDM	aspartate transami
40	809.5	39.6	449	2 T30955	probable aspartate
41	808	39.5	430	1 XNRTDM	aspartate transami
42	805	39.4	430	2 S35960	aspartate transami
43	802.5	39.2	414	2 T15494	aspartate transami
44	801	39.2	423	1 XNCHDM	aspartate transami
45	798.5	39.0	433	2 S01174	aspartate transami

ALIGNMENTS

RESULT 1
XNECD

aspartate transaminase (EC 2.6.1.1) aspC [validated] - Escherichia coli (strain K-12)
N:Alternate names: aspartate aminotransferase; transaminase A
C:Species: Escherichia coli
C:Date: 28-May-1986 #sequence revision 28-May-1986 #text_change 09-Jul-2004
C:Accession: A00598; A38045; A29306; G64832
R:Kuramitsu, S.; Okuno, S.; Ogawa, T.; Ogawa, H.; Kagamiyama, H.
J. Biochem. 97, 1259-1262, 1985
A:Title: Aspartate aminotransferase of Escherichia coli: nucleotide sequence of the aspC
A:Reference number: A00598; MUID:85289110; PMID:3897210
A:Accession: A00598
A:Molecule type: DNA
A:Residues: 1-396 <KUR>
A:Cross-references: UNIPROT:P00509; UNIPARC:UPI000011044B; GB:X05904; NID:g41012; PIDN:AA
R:Kondo, K.; Wakabayashi, S.; Yagi, T.; Kagamiyama, H.
Biochem. Biophys. Res. Commun. 122, 62-67, 1984
A:Title: The complete amino acid sequence of aspartate aminotransferase from Escherichia
A:Reference number: A38045; MUID:84256832; PMID:6378205
A:Accession: A38045
A:Molecule type: protein
A:Residues: 1-396 <KON1>
A:Cross-references: UNIPARC:UPI000011044B
R:Kondo, K.; Wakabayashi, S.; Kagamiyama, H.
J. Biol. Chem. 262, 8648-8659, 1987
A:Title: Structural studies on aspartate aminotransferase from Escherichia coli. Covalen
A:Reference number: A29306; MUID:87250482; PMID:3298240
A:Accession: A29306
A:Molecule type: protein
A:Residues: 1-396 <KON2>
A:Cross-references: UNIPARC:UPI000011044B
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64832
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <BLAT>
A:Cross-references: UNIPARC:UPI000011044B; GB:AE000195; GB:U00096; NID:g1787156; PIDN:AA
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: aspC
A:Map position: 21 min
C:Complex: homodimer
C:Function:
A:Description: catalyzes the reversible transfer of the amino group from L-aspartate to
A:Pathway: aspartate catabolism
A>Note: In eukaryotes, there are two isozymes: one is located in the mitochondrial matrix
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; homodimer; phosphoprotein; pyridoxal phosphate

F:246/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental
F:374/Binding site: substrate (Arg) #status predicted

Query Match 100.0%; Score 2045; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 9.3e-153;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETKTPTVLTSVKKAEOYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETKTPTVLTSVKKAEOYLLENE 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGCALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGCALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKELIV 240
Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKEQVLR 360
Qy 361 LREBFVGIVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREBFVGIVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 2

A85619
aspartate aminotransferase [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A85619
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <STO>
A:Cross-references: UNIPROT:Q8XDF3; UNIPARC:UPI00000D037B; GB:AE005174; NID:g12514100; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: aspC
C:Superfamily: aspartate aminotransferase

Query Match 99.6%; Score 2037; DB 2; Length 396;
Best Local Similarity 99.5%; Pred. No. 4e-152;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETKTPTVLTSVKKAEOYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETKTPTVLTSVKKAEOYLLENE 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGCALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGCALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

Qy 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKELIV 240
Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKEQVLR 360
Qy 361 LREBFVGIVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREBFVGIVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 3

C90755
aspartate aminotransferase [imported] - Escherichia coli (strain O157:H7, substrain RIMD
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: C90755
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90755
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <HAY>
A:Cross-references: UNIPROT:Q8XDF3; UNIPARC:UPI00000D037B; GB:BA000007; PIDN:BAB34434.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs1011
C:Superfamily: aspartate aminotransferase

Query Match 99.6%; Score 2037; DB 2; Length 396;
Best Local Similarity 99.5%; Pred. No. 4e-152;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETKTPTVLTSVKKAEOYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETKTPTVLTSVKKAEOYLLENE 60
Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGCALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTAQTPGGTGCALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Db 121 KRVVSNPSPWNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRAFAAMHKELIV 240
Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKEQVLR 360
Qy 361 LREBFVGIVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREBFVGIVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 4

AD0616

aspartate aminotransferase [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0616
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conner, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moulie, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0616
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <PAR>
A:Cross-references: UNIPARC:UP1000005A061; GB:AL513382; PIDN:CAD05398.1; PID:g16502160;
C:Genetics:
A:Gene: Sry1000
C:Superfamily: aspartate aminotransferase
Query Match 96.7%; Score 1978; DB 2; Length 396;
Best Local Similarity 95.7%; Pred. No. 1.7e-147;
Matches 379; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Qy 61 TTKNYLIGDIGIPEFCRCCTQELLFGKGSALINDKARTATQPGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLIGDIGIPEFCRCCTQELLFGKGSALINDKARTATQPGTGALRVAADFLAKNTSV 120
Qy 121 KRVSNSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 121 KRVSNSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAFGFARGLDEAGLRAFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAFGFARGLDEAGLRAFAAMHKLIV 240
Qy 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Db 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Qy 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Db 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Qy 361 LREERGVAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
Db 361 LREERGVAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
RESULT 5
AD0172
aspartate transaminase (EC 2.6.1.1) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD0172
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0172
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <KUR>
A:Cross-references: UNIPROT:Q8ZG95; UNIPARC:UP100000CD7E2; GB:AL590842; PIDN:CAC90239.1;
C:Genetics:
A:Gene: aspc

C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase
Query Match 86.8%; Score 1775; DB 2; Length 396;
Best Local Similarity 86.1%; Pred. No. 1.5e-131;
Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Qy 61 TTKNYLIGDIGIPEFCRCCTQELLFGKGSALINDKARTATQPGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLIGDIGIPEFCRCCTQELLFGKGSALINDKARTATQPGTGALRVAADFLAKNTSV 120
Qy 121 KRVSNSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 121 KRVSNSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAFGFARGLDEAGLRAFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWOTLAQSLVEKGMPLFDFAFGFARGLDEAGLRAFAAMHKLIV 240
Qy 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Db 241 ASSYKFNGLYNERVAGCTLVAADETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
Qy 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Db 301 DALRAIWQEQLTDMQRQIRQRMQLFVNTLQKGNRDFSFIKONGMFSFGLTKQVLR 360
Qy 361 LREERGVAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
Db 361 LREERGVAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
RESULT 6
I64132
aspartate transaminase (EC 2.6.1.1) - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: I64132
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <TIGR>
A:Cross-references: UNIPROT:P44425; UNIPARC:UP100001250E7; GB:U32835; GB:L42023; NID:g15
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:246/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 65.0%; Score 1329; DB 2; Length 396;
Best Local Similarity 62.6%; Pred. No. 1.5e-96;
Matches 248; Conservative 61; Mismatches 87; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKPTVLTTSVKAEQYLLNE 60
Qy 61 TTKNYLIGDIGIPEFCRCCTQELLFGKGSALINDKARTATQPGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLIGDIGIPEFCRCCTQELLFGKGSALINDKARTATQPGTGALRVAADFLAKNTSV 120
Qy 121 KRVSNSPWNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
Db 121 QNVWISTPTWPNHNAIFNAVGMTIREYRYDAERKALDWEHLEDSQASEGDDVLLHGC 180


```
Db 303 ELKGLWQBEVGEGRSRIASLRIGLVEALPHGLAERFAHVGAQRGMFSYTGLSPOQVARL 362
Qy 362 REEFGVYAVASGRVNVAGTDPDNMAPLCEAIVAV 395
Db 363 RDEHAVLVSSGRANVAGLDARRDLRLAQATAQV 396

RESULT 12
C82855
aromatic-amino-acid transaminase (EC 2.6.1.57) tyrB XF0036 [similarity] - Xylella fastidiosa
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82855
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82855
A:Status: preliminary
A:Genetics:
A:Molecule type: DNA
A:Residues: 1-400 <SIN>
A:Cross-references: UNIPROT:Q9PHA8; UNIPARC:UPI00000C229A; GB:AE003858; GB:AE003849; NID
A:Experimental source: strain 985C
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matuskuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshakano, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
A:Genetics:
A:Gene: XF0036
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase

Query Match 44.4%; Score 909; DB 2; Length 400;
Best Local Similarity 45.7%; Pred. No. 1.4e-63;
Matches 181; Conservative 76; Mismatches 137; Indels 2; Gaps 2;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIGYKDETGTPTVLTSVKKAQYLLNE 60
Db 3 LFTDVELVGPDPILSLNDYNTADTRTNKVNIGIGYCDSEGCIPLLRAVQQVEQLAKHP 62

Qy 61 TTKNYLGDIGIPEFCRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVADFLAKNTSV 120
Db 63 KPRGYLPIDGLPAYIKATQQLFLFGVDSPLLTAGRVATVSQTITGGSGALRVTAEVLLKQVLP 122

Qy 121 KRVVYNSPNWPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 123 ATVAISRFSWENHRALFTAAEFKTEDITYFTDLGHAVDFTGMVADLAKLQPTVVLLHGC 182

Qy 181 CHNPTGIDPTLEQWOTLAQLSVKGLWPLFPFAYQGFARGLEEDAEGLRAFAMH-KELI 239
Db 183 CHNPTGADLSRDQWQLVALFQERQLLPCIDLAYQFGNQGDADAYAIRLLAAEGISNVY 242

Qy 240 VASSYSKNGFLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
Db 243 VANSYSKFSYGERVGGLSIVASNTAQAIQSQVKRIIRTIYSSPSAHGAYILVAGVLN 302

Qy 300 NDALRAIWEQELTDMRQIRMRQLFVNTLQEKGANRDFSFIIKQNGMFSFSGLTKEQVL 359
Db 303 SHELTLWEQELTQWREIHGLRAGLVARLALTGA-PEFDFFIQRQAGNFSYSGLSKIQVD 361

Qy 360 RLREBEFGVYAVASGRVNVAGTDPDNMAPLCEAIVAV 395
```

```
Db 362 RLREBEFGVYAVSSGRICVAALSQHKLEYVAQAVVKV 397

RESULT 13
JC5124
aspartate transaminase (EC 2.6.1.1), cytosolic - rice
N:Alternate names: cytoplasmic aspartate aminotransferase; cytoplasmic glutamic-oxaloace
C:Species: Oryza sativa (rice)
C:Date: 02-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5124
R:Song, J.; Yanamoto, K.; Shomura, A.; Yano, M.; Minobe, Y.; Saeaki, T.
DNA Res. 3, 303-310, 1996
A:Title: Characterization and mapping of cDNA encoding aspartate aminotransferase in ric
A:Reference number: JC5124; MUID:97191542; PMID:9039500
A:Accession: JC5124
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-407 <SON>
A:Cross-references: UNIPROT:P37833; UNIPARC:UPI0000001116; GB:D14673; NID:g287297; PIDN:
A:Comment: This enzyme catalyzes the reversible transamination between a dicarboxylic am
C:Genetics:
A:Map position: 1
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:253/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 44.4%; Score 909; DB 2; Length 407;
Best Local Similarity 45.0%; Pred. No. 1.5e-63;
Matches 179; Conservative 69; Mismatches 148; Indels 2; Gaps 2;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIGYKDETGTPTVLTSVKKAQYLLNE 60
Db 6 VFAGLAQAPEDPILGVTVAYNKDPSPVKVNLGVGAYRTEGKPLVLNVVRAEQMLINNP 65

Qy 61 T-TKNYLGIDIGIPEFCRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVADFLAKNTS 119
Db 66 SRVKEYLPITGLADFNFKNLSAKLIFGADSPAIQENRVATVOCLSGTSGTLRVGGEFLARHYH 125

Qy 120 KRVVYNSPNWPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHG 179
Db 126 ERTIYIPOQTWGNHPKVFTLAGLTVRSYRYYPATRGDLDFQGLLEDLSAPSGAIVLLHA 185

Qy 180 CHNPTGIDPTLEQWOTLAQLSVKGLWPLFPFAYQGFARG-LEEDAEGLRAFAMH-KEL 238
Db 186 CAHNPTGVDPDPTLDQWEQIRQLMRSKALLPPFDSDAYQGFASGSLDQDAQSRMVFADGGEL 245

Qy 239 IVASSYSKNGFLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATIL 298
Db 246 LMAQSYAKNMGLYGERVGALSIIVCGSADVAVRVESQKLVIRPMYSNPPIHGASIVATIL 305

Qy 299 SNDALRAIWEQELTDMRQIRMRQLFVNTLQEKGANRDFSFIIKQNGMFSFSGLTKEQV 358
Db 306 KDSAMFNWETVELKGMADRIISMRRQLFDALKTRTETPGDWSHIIKQIGMFTTGLNSDQV 365

Qy 359 RLREBEFGVYAVASGRVNVAGTDPDNMAPLCEAIVAVL 396
Db 366 AFMRQEVHYWTSDDRISMGSLSGRTIPHLADAIHAAV 403

RESULT 14
S53303
aspartate transaminase (EC 2.6.1.1) AAT2 - proso millet
N:Alternate names: aspartate aminotransferase, cytosolic
C:Species: Panicum miliaceum (proso millet)
C:Date: 14-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S53303; S22378; S40067; S18891
R:Taniguchi, M.; Mori, J.; Sugiyama, T.
Plant Mol. Biol. 26, 723-734, 1994
A:Title: Structure of genes that encode isozymes of aspartate aminotransferase in Panicu
A:Reference number: S53303; MUID:95036052; PMID:7948926
A:Accession: S53303
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
```


A:Molecule type: DNA
A:Residues: 1-409 <TAN>
A:Cross-references: UNIPROT:Q42391; UNIPARC:UPI00000AA577; EMBL:D25322; NID:G435456; PID
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
R:Taniguchi, M.; Sawaki, H.; Sasakawa, H.; Hase, T.; Sugiyama, T.
Eur. J. Biochem. 204, 611-620, 1992
A:Title: Cloning and sequence analysis of cDNA encoding aspartate aminotransferase isozy
A:Reference number: S22377; MUID:92174917; PMID:1541276
A:Accession: S22378
A:Molecule type: mRNA
A:Residues: 1-409 <TAN>
A:Cross-references: UNIPARC:UPI00000AA577; EMBL:X63429; NID:G20598; PIDN:CAA45023.1; PID
A:Accession: S40067
A:Molecule type: protein
A:Residues: 72-84; 'X', 130-136; 213-230 <TAN2>
A:Cross-references: UNIPARC:UPI00001754F5; UNIPARC:UPI00001754F6; UNIPARC:UPI00001754F7
C:Genetics:
A:Gene: AAT2
A:Introns: 22/3; 46/3; 67/2; 95/2; 128/3; 169/1; 223/3; 269/3; 310/2; 342/1; 382/2
C:Superfamily: aspartate aminotransferase
C:Keywords: aspartate aminotransferase; cytosol; phosphoprotein; pyridoxal phosphate
F:255/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 44.4%; Score 908; DB 2; Length 409;
Best Local Similarity 45.5%; Pred. No. 1.8e-63;
Matches 181; Conservative 67; Mismatches 148; Indels 2; Gaps 2;
QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETGTPTVLTSVKKAQVLLNE 60
DB 8 VFAGIAQAPEDPILGVTVAFNKPDPSPVKNLGVGAYRTTEEGKPLVNLVVRRAEQMLNDP 67
QY 61 T-TKNYLGDIDPTEQRCCTOELLFGKGSALINDKARTAQTPGGTGCALRVAAADFLAKNTS 119
DB 68 SRVKEYLPITGLAEYNKLSAKLIFGADSPAIOENRVATVQCLSGTSLRGVGEFLAKHYH 127
QY 120 VKRVMVSNPSPNHSKVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLFHG 179
DB 128 ERTIVPVTWGNHPKVFTLAGLTVRSYRYYPATRGDLDFNGLLEDLSAPLGSLVLLHA 187
QY 180 CCHNPTGIDPTLEOMOTLAQLSVEKGLPLDFAYOGFARG-LEEDAELRAFAAMHKL 238
DB 188 CAHNPTGVDPDTEQEQIRQLMRSKSLPFFDSAYOGFASGLDADAQSVRFVADGGEL 247
QY 239 IVASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATIL 298
DB 248 LAAQSYAKNMGYGERVGSALVCSGADVAARVESQKLVIRPMYSSPPLHGPSVVATIL 307
QY 299 SNDALRAIWEQELTDMRQRIORMQLFVNTLOEKGANRDFPIIKONGMFSFSLTKQOV 358
DB 308 KDSEMFHEWTVLKMADRIISMROQLFDALKSRGTPGDWSHIIKQIGMFTFTGLNSEQV 367
QY 359 LRLREFGVYAVASGRVNVAGMTDPNMAPLCEAIVAVL 396
DB 368 AFMRQYHIYMTSDGRISMAGLNKKNVPHLADAIHAAV 405

RESULT 15
T14311
aspartate transaminase (EC 2.6.1.1), cytosolic [similarity] - carrot
N:Alternate names: aspartate aminotransferase
C:Species: Daucus carota (carrot)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14311
R:Turano, F.J.; Weisemann, J.M.; Matthews, B.F.
submitted to the EMBL Data Library, July 1992
A:Description: Identification and expression of a cDNA clone encoding aspartate aminotra
A:Reference number: Z17971
A:Accession: T14311
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-405 <TUR>
A:Cross-references: UNIPROT:P28734; UNIPARC:UPI00001250D4; EMBL:M92660; NID:G167545; PID
C:Function:

A:Description: reversibly catalyzes transamination between L-aspartate and 2-oxoglutarate
C:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:251/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 44.2%; Score 904; DB 2; Length 405;
Best Local Similarity 44.7%; Pred. No. 3.6e-63;
Matches 178; Conservative 67; Mismatches 151; Indels 2; Gaps 2;
QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETGTPTVLTSVKKAQVLLNE 60
DB 4 VFANVVRAPEDPILGVTVAYHKDQSPKLNGLGVAYRTTEEGKPLVNLVNVKAAEQMLVNDQ 63
QY 61 T-TKNYLGDIDPTEQRCCTOELLFGKGSALINDKARTAQTPGGTGCALRVAAADFLAKNTS 119
DB 64 SRVKEYLPITGLADPDKLSAKLIFGADSPAIOENRVATVQCLSGTSLRGVGEFLARHYH 123
QY 120 VKRVMVSNPSPNHSKVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAQGVVLFHG 179
DB 124 EHTVYIPQPTWGNHPKIFTLAGLSVKTYRYYPATRGDLDFEGMLLEDLSAPLGAIVLLHA 183
QY 180 CCHNPTGIDPTLEOMOTLAQLSVEKGLPLDFAYOGFARG-LEEDAELRAFAAMHKL 238
DB 184 CAHNPTGVDPDTEQEQIRQLMRSKSLPFFDSAYOGFASGLDADAQSVRFVADGGEC 243
QY 239 IVASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATIL 298
DB 244 LAAQSYAKNMGYGERVGSALVCSGADVAARVESQKLVIRPMYSSPPLHGCASIVAAIL 303
QY 299 SNDALRAIWEQELTDMRQRIORMQLFVNTLOEKGANRDFPIIKONGMFSFSLTKQOV 358
DB 304 KGDGLYNWTLLEKAMADRIISMROELFNALQAKGTPGDWSHIVKQIGMFTFTGLNSEQV 363
QY 359 LRLREFGVYAVASGRVNVAGMTDPNMAPLCEAIVAVL 396
DB 364 TFMNTVEYHIYLTSDGRISMAGLSRTVPHLADAIHAAV 401

Search completed: March 14, 2006, 18:00:45
Job time : 41 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 17:53:05 ; Search time 230 Seconds
(without alignments)
1214.736 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITAAPADPILGLDLF.....VAGMTPDNAPLCEAIVAVL 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2045	100.0	396	1	AAT ECOLI
2	2037	99.6	396	2	Q8XDF3 ECO57
3	2033	99.4	396	2	Q8FJ99 ECOL6
4	2026	99.1	396	2	Q8JLNA SHIFL
5	1982	96.9	396	1	AAT_SALT
6	1979	96.8	396	2	Q57R00 SALCH
7	1978	96.7	396	1	AAT_SALT
8	1978	96.7	396	2	Q5PGE8 SALPA
9	1780	87.0	396	2	Q6D451 ERWCT
10	1775	86.8	396	2	Q8ZG95 YERPE
11	1775	86.8	396	2	Q6CG9 YERPE
12	1775	86.8	401	2	Q8D050 YERPE
13	1682	82.2	396	2	Q7N625 PHOLL
14	1349	66.0	396	2	Q5D191 9PAST
15	1343	65.7	396	2	Q4QL54 HA8I8
16	1336	65.3	399	2	Q54R3 VIBF1
17	1333	65.2	396	2	Q6LPQ0 PHOPR
18	1329	65.0	396	1	AAT_HAEIN
19	1325	64.8	397	2	Q8EEM8 SHEON
20	1325	64.8	414	2	Q8YNG8 VIBPA
21	1324	64.7	396	2	Q9RNA0 AERHY
22	1318	64.4	399	2	Q8DAF6 VIBVU
23	1317	64.4	396	2	Q5F6U3 NEIG1
24	1316	64.4	399	2	Q7MJR4 VIBVU
25	1315	64.3	413	2	Q9KSG3 VIBCH
26	1311	64.1	397	2	Q9JVS3 NEIMA
27	1310	64.1	397	2	Q9KOP5 NEIMB
28	1299	63.5	396	2	Q9CN25 PASMU
29	1279	62.5	420	2	Q65TS2 MANSM
30	1257	61.5	396	2	Q7VLL0 HAEDU
31	1125.5	55.0	406	2	Q7VR08 CANBF

32	1081	52.9	397	2	Q7UHY8 RH0BA
33	1055	51.6	398	2	Q6AQM7 DESPS
34	1051.5	51.4	397	2	Q53137 MORSP
35	1044	51.1	196	2	Q5INQ0 ECOLI
36	1044	51.1	196	2	Q5INQ8 ECOLI
37	1044	51.1	196	2	Q5INS8 SHIDY
38	1040	50.9	196	2	Q5INT6 SHIBO
39	1039	50.8	195	2	Q5INR1 ECOLI
40	1039	50.8	195	2	Q5INS1 SHIDY
41	1038	50.8	195	2	Q5INS6 SHIDY
42	1038	50.8	195	2	Q5INT8 SHISO
43	1038	50.8	195	2	Q5INT9 SHIBO
44	1033	50.5	194	2	Q5IHH8 ECOLI
45	1033	50.5	194	2	Q5INS5 SHIDY

ALIGNMENTS

RESULT 1
AAT ECOLI
ID AAT ECOLI STANDARD; PRT; 396 AA.
AC P00509;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1) (Transaminase A) (ASPART).
GN Name=aspC; OrderedLocusNames=b0928;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=85289110; PubMed=3897210;
RA Kuramitsu S., Okuno S., Ogawa T., Ogawa H., Kagamiyama H.;
RT "Aspartate aminotransferase of Escherichia coli: nucleotide sequence of the aspC gene.";
RL J. Biochem. 97:1259-1262(1985).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86242111; PubMed=3521591;
RA Fotheringham I.G., Dacey S.A., Taylor P.P., Smith T.J., Hunter M.G., Finlay M.E., Primrose S.B., Parker D.M., Edwards R.M.;
RT "The cloning and sequence analysis of the aspC and tyrB genes from Escherichia coli K12. Comparison of the primary structures of the aspartate aminotransferase and aromatic aminotransferase of E. coli with those of the pig aspartate aminotransferase isoenzymes.";
RL Biochem. J. 234:593-604(1986).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1555;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kishimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [5]

```
RP PROTEIN SEQUENCE.
RX MEDLINE=84256832; PubMed=6378205;
RA Kondo K., Wakabayashi S., Yagi T., Kagamiyama H.;
RT "The complete amino acid sequence of aspartate aminotransferase from
RL Escherichia coli: sequence comparison with pig isoenzymes.";
RN Biochem. Biophys. Res. Commun. 122:62-67(1984).
RP PROTEIN SEQUENCE.
RX MEDLINE=87250482; PubMed=3298240;
RA Kondo K., Wakabayashi S., Kagamiyama H.;
RT "Structural studies on aspartate aminotransferase from Escherichia
RL coli. Covalent structure.";
RN J. Biol. Chem. 262:8648-8659(1987).
RP PROTEIN SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RL in the genome of Escherichia coli K-12.";
RN Electrophoresis 18:1259-1313(1997).
RP MUTAGENESIS OF TYR-65.
RX MEDLINE=91329346; PubMed=1868057;
RA Inoue K., Kuramitsu S., Okamoto A., Hirotsu K., Higuchi T.,
RA Kagamiyama H.;
RT "Site-directed mutagenesis of Escherichia coli aspartate
RL aminotransferase: role of Tyr70 in the catalytic processes.";
RN Biochemistry 30:7796-7801(1991).
RP MUTAGENESIS OF HIS-133.
RX MEDLINE=91177849; PubMed=2007566;
RA Yano T., Kuramitsu S., Tanase S., Morino Y., Hiromi K., Kagamiyama H.;
RT "The role of His143 in the catalytic mechanism of Escherichia coli
RL aspartate aminotransferase.";
RN J. Biol. Chem. 266:6079-6085(1991).
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF MUTANT ALA-246.
RX MEDLINE=90105323; PubMed=2513875;
RA Smith D.L., Almo S.C., Toney M.D., Ringe D.;
RT "2.8-A-resolution crystal structure of an active-site mutant of
RL aspartate aminotransferase from Escherichia coli.";
RN Biochemistry 28:8161-8177(1989).
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS), AND MUTAGENESIS OF ARG-374.
RX MEDLINE=91129283; PubMed=1593208;
RA Danishefsky A.T., Omufu J.J., Petsko G.A., Ringe D.;
RT "Activity and structure of the active-site mutants R386Y and R386F of
RL Escherichia coli aspartate aminotransferase.";
RN Biochemistry 30:1980-1985(1991).
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF MUTANT.
RX MEDLINE=99107891; PubMed=9891001; DOI=10.1074/jbc.274.4.2344;
RA Oue S., Okamoto A., Yano T., Kagamiyama H.;
RT "Redesigning the substrate specificity of an enzyme by cumulative
RL effects of the mutations of non-active site residues.";
RN J. Biol. Chem. 274:2344-2349(1999).
CC -|- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -|- COFACTOR: Pyridoxal phosphate.
CC -|- SUBUNIT: Homodimer.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the class-I pyridoxal-phosphate-dependent
CC aminotransferase family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; X03629; CAA2279.1; -; Genomic DNA.
DR EMBL; X05904; CAA29333.1; -; Genomic DNA.
```

KW 3D-structure; Amino transferase; Complete proteome;
 KW Direct protein sequencing; Pyridoxal phosphate; Transferase.

Query Match 100.0%; Score 2045; DB 1; Length 396;
 Best Local Similarity 100.0%; Pred. No. 3e-146;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETKPTVLTSVKKAEOYLLENE 60
 DB 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETKPTVLTSVKKAEOYLLENE 60

QY 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSV 120

QY 121 KRVMVSNPWNHKSVPNSAGLEVREYAYYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
 DB 121 KRVMVSNPWNHKSVPNSAGLEVREYAYYDAENHTLDFDALINSINEAQAGDVVLFHGC 180

QY 181 CHNPTGIDPTLQWOTLAQSVKQWPLPLDFAYQGFARGLDEDAEGLRAFAAMHKELIV 240
 DB 181 CHNPTGIDPTLQWOTLAQSVKQWPLPLDFAYQGFARGLDEDAEGLRAFAAMHKELIV 240

QY 241 ASSYKNGFLYNERVGACTLVAADSETVDRAFSONKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYKNGFLYNERVGACTLVAADSETVDRAFSONKAAIRANYSNPPAHGASVATILSN 300

QY 301 DALRAIWEQELTDMQRQIRORMQLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRORMQLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360

QY 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 2

QXDF3 ECO57
 ID Q8XDF3 ECO57 PRELIMINARY; PRT; 396 AA.
 AC Q8XDF3_07AG50;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Aspartate aminotransferase.
 GN Name=aspC; OrderedLocusNames=ECs1011, z1275;
 OS Escherichia coli O157:H7;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533 (2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=O157:H7 / Sakai / RMD 050952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12."
 RT DNA Res. 8:11-22 (2001).
 DR EMBL; AE005174; AAG55413.1; -; Genomic_DNA.

DR EMBL; BA000007; BAB34434.1; -; Genomic_DNA.
 DR PIR; A85619; A85619.
 DR PIR; C90755; C90755.
 DR HSSP; P00509; IART.
 DR SMR; Q8XDF3; 1-396.
 DR GO; GO:0008483; F:transaminase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR004839; Amino trans I/II.
 DR InterPro; IPR000796; Asp trans I/BS.
 DR InterPro; IPR004838; NH2transf I/BS.
 DR Pfam; PF00155; Amino tran 1_2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA TRANSFER CLASS 1; 1.
 KW Amino transferase; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43645 MW; 81C2063CE24DF08 CRC64;

Query Match 99.6%; Score 2037; DB 2; Length 396;
 Best Local Similarity 99.5%; Pred. No. 1.2e-145;
 Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETKPTVLTSVKKAEOYLLENE 60
 DB 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETKPTVLTSVKKAEOYLLENE 60

QY 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSV 120

QY 121 KRVMVSNPWNHKSVPNSAGLEVREYAYYDAENHTLDFDALINSINEAQAGDVVLFHGC 180
 DB 121 KRVMVSNPWNHKSVPNSAGLEVREYAYYDAENHTLDFDALINSINEAQAGDVVLFHGC 180

QY 181 CHNPTGIDPTLQWOTLAQSVKQWPLPLDFAYQGFARGLDEDAEGLRAFAAMHKELIV 240
 DB 181 CHNPTGIDPTLQWOTLAQSVKQWPLPLDFAYQGFARGLDEDAEGLRAFAAMHKELIV 240

QY 241 ASSYKNGFLYNERVGACTLVAADSETVDRAFSONKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYKNGFLYNERVGACTLVAADSETVDRAFSONKAAIRANYSNPPAHGASVATILSN 300

QY 301 DALRAIWEQELTDMQRQIRORMQLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRORMQLFVNTLOEGKANRDFSFIKONGMFSGLTKEQVLR 360

QY 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 3
 Q8FJ99 ECO6
 ID Q8FJ99 ECO6 PRELIMINARY; PRT; 396 AA.
 AC Q8FJ99;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Aspartate aminotransferase (EC 2.6.1.1).
 GN Name=aspC; OrderedLocusNames=c1070;
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=O6:H1 / CPT073 / ATCC 700928 / UPEC;
 RX MEDLINE=2338234; PubMed=12471157; DOI=10.1073/pnas.252529799;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesech P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016758; AAN79538.1; -; Genomic_DNA.
DR HSSP; P00509; IART.
DR SMR; Q8FJ99; 1-396.
DR GO; GO:0004069; P:aspartate transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.
DR InterPro; IPR00796; AspTrans.
DR InterPro; IPR004838; NhtTransf_1_BS.
DR Pfam; PF00155; AminoTran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Complete proteome.
SQ SEQUENCE 396 AA; 43568 MW; 9FA46ECB413FC679 CRC64;

Query Match 99.4%; Score 2033; DB 2; Length 396;
Best Local Similarity 99.0%; Pred. No. 2.4e-145;
Matches 392; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPVLTSVKKAQVYLENE 60
DB 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPVLTSVKKAQVYLENE 60

QY 61 TTKNYLGIDGIDPEFCRCRQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
DB 61 TTKNYLGIDGIDPEFCRCRQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120

QY 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

QY 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRFAAAMHKELIV 240
DB 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRFAAAMHKELIV 240

QY 241 ASSYSKNFGLYNERVAGCTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300
DB 241 ASSYSKNFGLYNERVAGCTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300

QY 301 DALRAIWEQELTDMRQRLQRMQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360
DB 301 DALRAIWEQELTDMRQRLQRMQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360

QY 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
DB 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 4
ID Q83LN4 SHIFL PRELIMINARY; PRT; 396 AA.
AC Q83LN4_07C286;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Aspartate aminotransferase.
GN Name=aspC; OrderedLocusNames=S0989; SF0925;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2227406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
```

```

RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE005674; AAN42554.1; -; Genomic_DNA.
DR EMBL; AE016981; AAP16440.1; -; Genomic_DNA.
DR HSSP; P00509; IART.
DR SMR; Q83LN4; 1-396.
DR GO; GO:0008483; P:transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.
DR InterPro; IPR00796; AspTrans.
DR InterPro; IPR004838; NhtTransf_1_BS.
DR Pfam; PF00155; AminoTran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Amino transferase; Complete proteome; Transferase.
SQ SEQUENCE 396 AA; 43617 MW; 621C1FB9FBA4DD5A CRC64;

Query Match 99.1%; Score 2026; DB 2; Length 396;
Best Local Similarity 99.2%; Pred. No. 8.2e-145;
Matches 393; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPVLTSVKKAQVYLENE 60
DB 1 MFENITAAPADPILGLADLFRADERPGKINLGIGYKDETKTPVLTSVKKAQVYLENE 60

QY 61 TTKNYLGIDGIDPEFCRCRQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
DB 61 TTKNYLGIDGIDPEFCRCRQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120

QY 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 121 KRVMVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180

QY 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRFAAAMHKELIV 240
DB 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLFDFAVQGFARGLEEDAEGLRFAAAMHKELIV 240

QY 241 ASSYSKNFGLYNERVAGCTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300
DB 241 ASSYSKNFGLYNERVAGCTLVAADSETVDRAFSQMKAIRANYSNPPAHGASVVATILSN 300

QY 301 DALRAIWEQELTDMRQRLQRMQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360
DB 301 DALRAIWEQELTDMRQRLQRMQLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360

QY 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
DB 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 5
ID AAT_SALTY STANDARD; PRT; 396 AA.
AC P58661;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1) (Transaminase A) (AspAT).
GN Name=aspC; OrderedLocusNames=STM0998;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
```

OC Enterobacteriaceae; Salmonella.
 RN NCBI_TaxID=602;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Speth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 CC -!- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
 CC L-glutamate.
 CC -!- COPACTOR: Pyridoxal phosphate (By similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the class-I pyridoxal-phosphate-dependent
 CC aminotransferase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AE008743; AAL19932.1; -; Genomic_DNA.
 DR HSSP; P00509; IART.
 DR SMR; P58661; 1-396.
 DR StyGene; SG77777; aspC.
 DR InterPro; IPR004839; Aminotrans_I/II.
 DR InterPro; IPR000796; Asp_trans.
 DR InterPro; IPR004838; Nhrtransf_BS.
 DR PANTHER; PTHR11879; Asp_trans; 1.
 DR Pfam; PF00155; Aminotran 1.2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR KW Aminotransferase; Complete proteome; Pyridoxal phosphate; Transferase.
 FT ACT SITE 374 374
 FT BINDING 246 246 Pyridoxal phosphate (covalent) (By
 FT similarity).
 FT SQ SEQUENCE 396 AA; 43521 MW; 960940D0148D02FB CRC64;
 Query Match 96.9%; Score 1982; DB 1; Length 396;
 Best Local Similarity 96.0%; Pred. No. 1.8e-141;
 Matches 380; Conservative 11; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKAEQYLLNE 60
 DB 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKAEQYLLNE 60
 QY 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 QY 121 KRVSNSPNHKSQVFNAGLEVEYAYDAENHTLDPDALINSLEAQAGDVVLFHGC 180
 DB 121 KRVSNSPNHKSQVFNAGLEVEYAYDAENHTLDPDALINSLEAQAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWQTLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKELIV 240
 DB 181 CHNPTGIDPTLEQWQTLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKELIV 240
 QY 241 ASSYSKNGFLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYSKNGFLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMQRLFVNTLQEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMQRLFVNTLQEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 QY 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
 QY 241 ASSYSKNGFLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYSKNGFLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMQRLFVNTLQEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMQRLFVNTLQEGKANRDFSFIKQNGMFSFGLTKQVLR 360

QY 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
 RESULT 6
 Q57R00_SALCH PRELIMINARY; PRT; 396 AA.
 AC Q57R00;
 DT 10-MAY-2005 (TREMELrel. 30, Created)
 DT 10-MAY-2005 (TREMELrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMELrel. 30, Last annotation update)
 DE Aspartate aminotransferase.
 GN Name=aspC; OrderedLocusNames=SC0955;
 OS *Salmonella choleraesuis* (Salmonella enterica).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=591;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SC-B67;
 RX PubMed=15781495;
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
 RA Wang H.-S., Lee Y.-S.;
 RT "The genome sequence of *Salmonella enterica* serovar Choleraesuis, a
 RT highly invasive and resistant zoonotic pathogen.";
 RL Nucleic Acids Res. 33:1690-1698 (2005).
 DR EMBL; AE017220; AAX64861.1; -; Genomic DNA.
 KW Aminotransferase; Complete proteome; Transferase.
 SQ SEQUENCE 396 AA; 43537 MW; 891640CF08B02FB CRC64;
 Query Match 96.8%; Score 1979; DB 2; Length 396;
 Best Local Similarity 95.7%; Pred. No. 3e-141;
 Matches 379; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKAEQYLLNE 60
 DB 1 MFENITAAPADPILGLADLFRADRRPGKINLGIVGVDGTGTPVLTSTVKAEQYLLNE 60
 QY 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 DB 61 TTKNYLGIDGIPFGRCCTOELLFGKGSALINDKARTATPGGTGALRVAADFLAKNTSV 120
 QY 121 KRVSNSPNHKSQVFNAGLEVEYAYDAENHTLDPDALINSLEAQAGDVVLFHGC 180
 DB 121 KRVSNSPNHKSQVFNAGLEVEYAYDAENHTLDPDALINSLEAQAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQWQTLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKELIV 240
 DB 181 CHNPTGIDPTLEQWQTLAQLSVKQWLPFLDFAYOGFARGLEDAEGLRFAAMHKELIV 240
 QY 241 ASSYSKNGFLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 DB 241 ASSYSKNGFLYNERVAGCTLVAAADSETVDRAFQSKAAIRANYSNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMQRQIRMQRLFVNTLQEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 DB 301 DALRAIWEQELTDMQRQIRMQRLFVNTLQEGKANRDFSFIKQNGMFSFGLTKQVLR 360
 QY 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
 DB 361 LREEFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
 RESULT 7
 AAT_SALTI STANDARD; PRT; 396 AA.
 AC Q56114;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Aspartate aminotransferase (EC 2.6.1.1) (Transaminase A) (AspAT).
 GN Name=aspC; OrderedLocusNames=STY1000, t1936;

OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CT18;
 RA MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Parikhil J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
 RA Krogh A., Larsen T.S., Leach S., Mouton S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RX DOI=10.1128/JB.185.7.2330-2337.2003;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 1-74.
 RC STRAIN=IMSS-1;
 RA Fernandez-Mora M., Calva E.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
 CC L-glutamate.
 CC -|- SUBUNIT: Pyridoxal phosphate (By similarity).
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- SIMILARITY: Belongs to the class-I pyridoxal-phosphate-dependent
 CC aminotransferase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; AL627268; CAD05398.1; -; Genomic DNA.
 DR EMBL; AE016840; AAO69551.1; -; Genomic DNA.
 DR EMBL; X89757; CAA61906.1; -; Genomic DNA.
 DR HSSP; P00509; 1ART.
 DR SMR; Q56114; 1-396.
 DR InterPro; IPR004839; Aminotrans_I/II.
 DR InterPro; IPR000796; Asp_trans.
 DR InterPro; IPR004838; Nitransf_1_BS.
 DR PANTHER; PTHR11879; Asp_trans; 1.
 DR Pfam; PF00155; Aminotran_1.2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR Amino transferase; Complete proteome; Pyridoxal phosphate; Transferase.
 FT ACT SITE 374 374
 FT BINDING 246 246 Pyridoxal phosphate (covalent) (By
 FT similarity).
 FT SEQUENCE 396 AA; 43507 MW; 974C1585438D02FB CRC64;
 Query Match 96.7%; Score 1978; DB 1; Length 396;
 Best Local Similarity 95.7%; Pred. No. 3.5e-141;
 Matches 379; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEOYLLENE 60
 DB 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEOYLLENE 60

QY 61 TTKNYLIGDIGIPERCRCTOELLFGKGSALINDKXARTAOPTGGTCALRVAADFLAKNTSV 120
 DB 61 TTKNYLIGDIGIPEFARCTOELLFGKGSALINDKXARTAOPTGGTCALRIAADFLAKNTPV 120
 QY 121 KRVVSNPSPNKHKSFNVSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
 DB 121 KRVVSNPSPNKHKSFNVAAGLEVREYAYDAENHSLDFEALQASLSAQAGDVVLFHGC 180
 QY 181 CHNPTGIDPTLEQMOTLAQLSVEKGLPLDFAYOGFARGLEEDAELRAFAAMHKEIV 240
 DB 181 CHNPTGIDPTLEQMVAELSVKGLPLDFAYOGFARGLEEDAELRAFAALHKEIV 240
 QY 241 ASSYKNGFLYNERVAGCTIVAADSETVDRAFQSKMAAIRANYNPPAHGASVATILSN 300
 DB 241 ASSYKNGFLYNERVAGCTIVAADAETVDRAFQSKMAIRANYNPPAHGASVATILSN 300
 QY 301 DALRAIWEQELTDMRQIORMRQLFVNTLOEKGNRDFSFIIKONGMPSFGLTKEQVLR 360
 DB 301 DALRAIWEQELTDMRQIORMRQLFVNTLOEKGNRDFSFIIKONGMPSFGLTKDOVLR 360
 QY 361 LREEFGYVAVASGRVNVAGTTPDNMAPLCEAIVAVL 396
 DB 361 LREEFGYVAVASGRVNVAGTTPDNMAPLCEAIVAVL 396

RESULT 8

QSPG8 SALPA
 ID QSPG8 SALPA PRELIMINARY; PRT; 396 AA.
 AC QSPG8;
 DT 01-FEB-2005 (TREMBlrel. 29, Created)
 DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE Aspartate aminotransferase.
 GN Name=aspC; OrderedLocusNames=SPAI1800;
 OS Salmonella paratyphi-a
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=54388;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 9150;
 RX PubMed=15531882; DOI=10.1038/ng1470;
 RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
 RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
 RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
 RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
 RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
 RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
 RA Spieth J., Wilson R.K.;
 RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
 RT restricted serovars of Salmonella enterica that cause typhoid.";
 RL Nat. Genet. 36:1268-1274(2004).
 DR EMBL; CP000026; AAV7716.1; -; Genomic_DNA.
 DR SMR; QSPG8; 1-396.
 DR GO; GO:0008483; F:transaminase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR004839; Aminotrans_I/II.
 DR InterPro; IPR000796; Asp_trans.
 DR InterPro; IPR004838; Nitransf_1_BS.
 DR Pfam; PF00155; Aminotran_1.2; 1.
 DR PRINTS; PR00799; TRANSAMINASE.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR Amino transferase; Complete proteome; Transferase.
 KW SEQUENCE 396 AA; 43549 MW; 88E9AEDEEA9D149C CRC64;

Query Match 96.7%; Score 1978; DB 2; Length 396;
 Best Local Similarity 95.7%; Pred. No. 3.5e-141;
 Matches 379; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEOYLLENE 60
 DB 1 MFENITAPADPILGLADLFRADRPCKINIGVYKDETKPTVLTSVKKAEOYLLENE 60

1	MFENITAAPADPILGLADLPRADDRPGKINLIGVYKDETGKTPVLTSSVKKAEQYILLENE	60
61	TTKNYLIGDIGIPEFCRCTQELLFGKGSALINDKRRARTAQTPGTCGALRVAADFLAKNTSV	120
61	TTKNYLIGDIGIPEFARCTQELLFGKGSALINDKRRARTAQTPGTCGALRIAADFLAKNTPV	120
121	KRVWVSNPSWPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSLENAQAGDVVLFGCC	180
121	KRVWVSNPSWPNHKSVPNSAAGLEVREYAYDAENHTLDFEALQASLSEAQAGDVVLFGCC	180
181	CHNPTGIDPTLEQOMTTLAQLSVKQWLPFLFPAYQGFARGLEEDAEGRAPAAHKEHLIV	240
181	CHNPTGIDPTLEQOMQVLAELSVEKGWLPFLFPAYQGFARGLEEDAEGRAPFAALHKEHLIV	240
241	ASSYSKNFGLYNERVGACTLVAADSETVDRAFQOMKAARANYSNPPAHGASVATILSN	300
241	ASSYSKNFGLYNERVGACTLVAADAETVDRAFQOMKSAIRANYSNPPAHGASIVATILSN	300
301	DALRAIWEQEQLTDMQRQIORMROLFVNLTQEBKANRDFSFIKQNGMFSFGSLTKQEVLR	360
301	DVLRAIWEQEQLTDMQRQIORMROLFVNLTQEBKANRDFSFIKQNGMFSFGSLTKDQVLR	360
361	LRBEFGYVAVASGRVNVAGMTPDNNAPLCEAIVAVL	396
361	LRBEFGYVAVASGRVNVAGMTPDNNAPLCEAIVAVL	396

```

RESULT 9
Q6D451_ERWCT
ID Q6D451_ERWCT PRELIMINARY; PRT; 396 AA.
AC Q6D451;
DT 28-OCT-2004 (TrEMBLrel. 28, Created)
DT 28-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 28-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1).
GN Name=aspC; OrderedLocusNames=ECA5243;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
[1]
NN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Akin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX50851; CAG75442.1; -; Genomic_DNA.
DR SMR; Q6D451; 1-396.
DR GO; GO:0004069; F:aspartate transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/I.
DR InterPro; IPR000796; Asp_trans.
DR InterPro; IPR004838; NHTransf_1_BS.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
SQ Amino transferase; Complete proteome; Translate transferase.
KW SEQUENCE 396 AA; 43081 MW; 817240F2834105A7 CRC64;

```

Query Match 87.0%; Score 1780; DB 2; Length 396;
Best Local Similarity .84.6%; Pred. No. 3.4e-126;
Matches 335; Conservative 29; Mismatches 32; Indels

1	MFENISAAPADPILGLTDLFRADRAKINLIGIYVDETKTPVLT	SVKKAHYLLNE	60
61	TTKNYLGIIDGIPFGRCRTOELLFGKGSALINDKRAARTAO	TGGTGALRVAADFLAKNTSV	120
61	TTKNYLGIIDGLPAFGQCTOELLFGKQNAIIADKRAARTAO	TGGTGALRVAADFIANQTS	120
121	KRWVSNPSWPNHKSVPNSAGLEVREYAYDAENHTLDFDAL	INSLNEAQAGDVVLPHGC	180
121	KRIWISNPTWPNHNVSFSAGLEVCQDYDDAANHALDFDGL	NLSINAVEAGDVVLPHGC	180
181	CHNPTGIDPTLEOQTTLAQLSVKXGWLPLDFPAYQGFARG	LEEDAEGLRAFAAMHEKELIV	240
181	CHNPTGIDPTAEQWATLAELSVAKGWLPLDFPAYQGFARG	IEDEAEGRLPAAKHAELIV	240
241	ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQWKAARI	RANYSNPPHAGASVATILSN	300
241	CSSYSKNFGLYNERVGACTLVAADAAATADKAFSQVKAAI	RANYSNPPHSGATIVATILGN	300
301	DALRAIWEQELTDMRQRIQRMROLFVNTLQEBKANRDFSF	TIKQNGMFSFSGLTKEQOVL	360
301	DALKAIWEQELTAMREIRQMRQLFVNTLQEBKAQDFSF	TIKQNGMFSFSGLTKEQOVL	360
361	LRBEFGYVAVASGRVNVAGMTPDNNAPLCEAIVAVL		396
361	LRBEFGYVAVNSGRVNVAGMTPENNAPLCEAIVAVL		396

RESULT 10
Q8ZG95 YERPE PRELIMINARY; PRT; 396 AA.
ID Q8ZG95 YERPE PRELIMINARY; PRT; 396 AA.
AC Q8ZG95;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1).
GN Name=aspC; Ordered locus Names=yPO1410;
OS Yersinia pestis.
OC Bacteria, Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
NCBI_TaxID=632;
[1]
Nucleotide Sequence.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Itball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Chillingworth T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oysen P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR ENBL: AJ414148; CAC90239.1; -; Genomic DNA.

SNR; Q82G95; 1-396.
GO; GO:0004069; F.aspartate transaminase activity; IEA.
GO; GO:0006520; P.amino acid metabolism; IEA.
GO; GO:0009058; P.biosynthesis; IEA.
GO; GO:0004839; A.minotrans I/II.
InterPro; IPR0040796; Asp trans.
InterPro; IPR004838; NHtransf 1 BS.
Pfam; PF00155; Amino tran 1.2-1.
PRINTS; PR00799; TRANSAMINASE.
PROSITE; PS00105; AA TRANSFER CLASS 1; 1.
Aminotransferase; Complete proteome; Transferase.
SEQUENCE 396 AA; 43118 MW; F679902DC3EE9256 CRC64;
SEQUENCE

Query Match 86.8%; Score 1775; DB 2; Length 396;
Best Local Similarity 86.1%; Pred. NO. 8.1e-126;
Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

```

Qy 1 MFENITAAPADPILGLADLFRADRPCKINIGVYKDETCKTPTVLSVKKAEQVLLNE 60
Db 1 MFEKITAAPADPILGLTDIFRADDRAHKINILGIGYKDETKTPTVLSVKKAEQVLLNE 60
Qy 61 TTKNYLIGDGIPEFGRCCTQELLFGKGSALINDKRAATAQTGGTCALRVAADFLAKNTSV 120
Db 61 ATKNYLIGDGLPVFASCTQELLFGANSALIIADKRAATAQTGGTGGLRIADFAHQTS 120
Qy 121 KRVVSNPSPNHNKSVFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Db 121 KRVVSNPSPNHNKSVFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRAFAAMHKEIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRIFAATHQELIV 240
Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 CSSYSKNFGLYNERVGACTLVAADSNVADTAFSQKAVIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIQRMRLQFVNTLOEKGANRDFSFIIKQNGMFSFSLTKKEQVLR 360
Db 301 AALRAIWEQELTDMRQRIQRMRLQFVNTLOEKGAAQDFSFIIQNGMFSFSLTKKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LRDEFAYVAVNSGRVNVAGMTPDNMAPLCEAIVAVL 396

```

RESULT 11

```

ID Q66CG9 YERPS PRELIMINARY; PRT; 396 AA.
AC Q66CG9
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Aspartate aminotransferase (EC 2.6.1.1).
GN Name=aspC; OrderedLocusNames=YPTB1434;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype 1;
RX PubMed=1535858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH20674.1; -; Genomic_DNA.
DR SMR; Q66CG9; 1-396.
DR GO; GO:0004069; F:aspartate transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.
DR InterPro; IPR000796; AspTrans.
DR InterPro; IPR004838; NhrTrans_1_BS.
DR Pfam; PF00155; AminoTrans_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW AminoTransferase; Complete proteome; Transference.
SQ SEQUENCE 396 AA; 43118 MW; F679902DC5E9256 CRC64;

```

Query Match 86.8%; Score 1775; DB 2; Length 396;
 Best Local Similarity 86.1%; Pred. NO. 8.1e-126;
 Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

```

Qy 1 MFENITAAPADPILGLADLFRADRPCKINIGVYKDETCKTPTVLSVKKAEQVLLNE 60
Db 1 MFEKITAAPADPILGLTDIFRADDRAHKINILGIGYKDETKTPTVLSVKKAEQVLLNE 60
Qy 61 TTKNYLIGDGIPEFGRCCTQELLFGKGSALINDKRAATAQTGGTCALRVAADFLAKNTSV 120
Db 61 ATKNYLIGDGLPVFASCTQELLFGANSALIIADKRAATAQTGGTGGLRIADFAHQTS 120
Qy 121 KRVVSNPSPNHNKSVFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Db 121 KRVVSNPSPNHNKSVFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRAFAAMHKEIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRIFAATHQELIV 240
Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 CSSYSKNFGLYNERVGACTLVAADSNVADTAFSQKAVIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIQRMRLQFVNTLOEKGANRDFSFIIKQNGMFSFSLTKKEQVLR 360
Db 301 AALRAIWEQELTDMRQRIQRMRLQFVNTLOEKGAAQDFSFIIQNGMFSFSLTKKEQVLR 360
Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LRDEFAYVAVNSGRVNVAGMTPDNMAPLCEAIVAVL 396

```

RESULT 12

```

ID Q8D050 YERPE PRELIMINARY; PRT; 401 AA.
AC Q8D050; 074VU3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Aspartate aminotransferase.
GN Name=aspC; OrderedLocusNames=YPI183, Y2760;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song X., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans.";
RL DNA Res. 11:179-197(2004).
DR EMBL; AE013879; AAM86312.1; -; Genomic_DNA.
DR EMBL; AE017131; AAS61426.1; -; Genomic_DNA.
DR HSSP; P00509; 1ART.
DR SMR; Q8D050; 6-401.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.

```

```

DR InterPro; IPR000796; Asp_trans.
DR InterPro; IPR004838; NHTtransf_1_BS.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Aminotransferase; Transferase.
SQ SEQUENCE 401 AA; 43652 MW; A59AEBB191041831 CRC64;

Query Match      86.8%; Score 1775; DB 2; Length 401;
Best Local Similarity 86.1%; Pred. No. 8.3e-126;
Matches 341; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETKTPVLTSVKKAEQVLLNE 60
DB 6 MFEKITAAPADPILGLTDFRADDRAHKINLIGVYKDETKTPVLTSVKKAEQVLLNE 65
QY 61 TTKNYLGIDGIPFGRCCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
DB 66 ATKNYLGIDGLPVFASCTQELLFGANSALIIADKARTAQTPGGTGGLRIRAADFIAHQ TSA 125
QY 121 KRVVSNPSWPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 126 KRVVSNPSWPNHKNVFAAGLEVVEYAYDAENHALDFDGLLSLSAQAGDVVLFHGC 185
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGLRAFAAMHKE LIV 240
DB 186 CHNPTGIDPTETQWSQLAELSAKGLWLPFLDFAYOGFANGLEEDAQGLRIFAATHQELIV 245
QY 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQKAAIRANYSNPPAHGASVATILSN 300
DB 246 CSSYSKNFGLYNERVGACTLVAADSNVADTAFSQVKAIRANYSNPPAHGASVATILSN 305
QY 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQKGANRDFSFIIKONGMFSFGLTKEQVLR 360
DB 306 AALRAIWEQELTDMRQRIQRMQLFVNTLQKGAQDFSFIIKONGMFSFGLTKEQVLR 365
QY 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
DB 366 LRDEFVAVVNSGRVNVAGMTPDNNAPLCEAIVAVL 401

RESULT 13
Q7N625 PHOLL PRELIMINARY; PRT; 396 AA.
AC Q7N625
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Aspartate aminotransferase (Transaminase A).
GN Name=aspC; OrderedLocusNames=plu1750;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Glavaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dagaia B., Detose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siquier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens."
RL EMBL; BX571864; CAE14043.1; -; Genomic_DNA.
DR HSSP; P00509; 1AFT.
DR SMR; Q7N625; 1-396.
DR PhotocList; plu1750; -.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.

DR InterPro; IPR000796; Asp_trans.
DR InterPro; IPR004838; NHTtransf_1_BS.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Aminotransferase; Transferase.
SQ SEQUENCE 396 AA; 43575 MW; 7EC3B64A6DB6C26 CRC64;

Query Match      82.2%; Score 1682; DB 2; Length 396;
Best Local Similarity 79.8%; Pred. No. 8.9e-119;
Matches 316; Conservative 38; Mismatches 42; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETKTPVLTSVKKAEQVLLNE 60
DB 1 MFEKITAAPADPILGLADSFPSDPRNTKINLIGVYKDETKTPVLTSVKKAEQVLLNE 60
QY 61 TTKNYLGIDGIPFGRCCTQELLFGKGSALINDKARTAQTPGGTGALRVAADFLAKNTSV 120
DB 61 TTKNYLFISGLAEFRVTVQELLFGKHRTAQSPGGTGALRIRAADFIAKQNA 120
QY 121 KRVVSNPSWPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180
DB 121 KRVVSNPTWPNHKNVFAAGLEVVEYAYDAEKHALNFEDMLASLSAQAGDVVLFHGC 180
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGLRAFAAMHKE LIV 240
DB 181 CHNPTGIDPTPAQWAKLAEMSAEKGWLPFLDFAYOGFAPKGLNEAELRIFAKNHNLIV 240
QY 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQKAAIRANYSNPPAHGASVATILSN 300
DB 241 ASSYSKNFGLYNERVGACTIVASDSTAFAKAFSAQAKAIRANYSNPPAHGASIVTILSN 300
QY 301 DALRAIWEQELTDMRQRIQRMQLFVNTLQKGANRDFSFIIKONGMFSFGLTKEQVLR 360
DB 301 EDLKAWEQELTDMRQRIQRMQLFVNTLQKGAQDFSFIIKONGMFSFGLTKEQVLR 360
QY 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
DB 361 LRDEFVYAVSSGRVNVAGMTPDNNAPLCEAIVAVL 396

RESULT 14
Q5DI91_9PAST PRELIMINARY; PRT; 396 AA.
AC Q5DI91_9PAST
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE AspC.
GN Name=aspC;
OS Actinobacillus porcitonisillarum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=189834;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CCUG 46996;
RX PubMed=15863281; DOI=10.1016/j.vetmic.2005.01.020;
RA Kuhnert P., Schlatter Y., Frey J.;
RT "Characterization of the type I secretion system of the RTX toxin
RT AptXII in 'Actinobacillus porcitonisillarum'.";
RL Vet. Microbiol. 107:225-232 (2005).
DR EMBL; AY956600; AAX21413.1; -; Genomic_DNA.
DR InterPro; IPR004839; Aminotrans_1/II.
DR InterPro; IPR000796; Asp_trans.
DR Pfam; PF00155; Aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
SQ SEQUENCE 396 AA; 43374 MW; 605EA13BFCDF4BD0 CRC64;

Query Match      66.0%; Score 1349; DB 2; Length 396;

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 18:00:15 ; Search time 47 Seconds
(without alignments)
696.587 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score: 2045

Sequence: 1 MFENITAAPADPILGLADLF.....VAGWTPDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PGTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2045	100.0	396	2	US-08-985-908-24
2	1931	94.4	402	2	US-09-489-039A-12073
3	982	48.0	413	2	US-09-252-991A-28999
4	909.5	44.5	543	2	US-09-252-991A-29078
5	908	44.4	233	2	US-09-543-681A-7977
6	871.5	42.6	402	2	US-09-540-236-2235
7	844.5	41.3	445	2	US-09-489-039A-8472
8	843.5	41.2	419	2	US-09-543-681A-7295
9	830.5	40.6	427	2	US-09-328-352-4895
10	794	38.8	430	2	US-09-976-594-430
11	756	37.0	416	2	US-09-949-016-7087
12	726.5	35.5	395	2	US-09-198-452A-790
13	726.5	35.5	397	2	US-09-438-185A-742
14	694	33.9	365	2	US-09-801-874-5
15	542	26.5	303	2	US-09-248-796A-17477
16	486.5	23.8	289	2	US-09-248-796A-17478
17	436.5	21.3	141	2	US-09-543-681A-7889
18	359.5	17.6	380	2	US-09-801-874-2
19	321.5	15.7	241	2	US-09-248-796A-17480
20	284.5	13.9	188	2	US-09-248-796A-17479
21	233.5	11.4	254	2	US-09-801-874-4
22	191.5	9.4	123	2	US-09-513-999C-6182
23	182	8.9	115	2	US-09-248-796A-17476
24	161.5	7.9	434	2	US-09-134-001C-4695
25	153.5	7.5	429	2	US-09-710-279-3174
26	148	7.2	29	1	US-07-952-817-20
27	148	7.2	29	6	5210025-13

28	121.5	5.9	373	2	US-09-902-540-10929	Sequence 10929, A
29	121.5	5.9	771	2	US-09-712-363-157	Sequence 157, App
30	118.5	5.8	413	2	US-09-489-039A-13270	Sequence 13270, A
31	117	5.7	401	2	US-09-252-991A-27198	Sequence 27198, A
32	110.5	5.4	345	2	US-09-724-623-65	Sequence 65, Appl
33	109.5	5.4	378	2	US-09-489-039A-10110	Sequence 10110, A
34	108.5	5.3	393	2	US-09-543-681A-4673	Sequence 4673, Ap
35	108	5.3	843	1	US-09-172-977-3	Sequence 3, Appli
36	108	5.3	843	2	US-09-404-108-3	Sequence 3, Appli
37	107	5.2	1839	1	US-09-172-977-4	Sequence 4, Appli
38	107	5.2	1839	2	US-09-404-108-4	Sequence 4, Appli
39	107	5.2	2753	2	US-09-949-016-7659	Sequence 7659, Ap
40	107	5.2	2753	2	US-09-949-016-7660	Sequence 7660, Ap
41	107	5.2	3924	2	US-09-538-092-1246	Sequence 1246, Ap
42	106.5	5.2	422	2	US-09-252-991A-21571	Sequence 21571, A
43	102	5.0	395	2	US-09-489-039A-14008	Sequence 14008, A
44	102	5.0	551	2	US-09-538-092-738	Sequence 738, App
45	101.5	5.0	889	2	US-09-489-039A-7241	Sequence 7241, Ap

ALIGNMENTS

RESULT 1
US-08-985-908-24
; Sequence 24, Application US/08985908
; Patent No. 6004773
; GENERAL INFORMATION:
; APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI N.
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JERPERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/985,908
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-325659
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-985-908-24

Query Match 100.0%; Score 2045; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.2e-205;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADPGKINLIGVYKDETKPVLTSVKAEQYLLENE 60
|||||
Db 1 MFENITAAPADPILGLADLFRADPGKINLIGVYKDETKPVLTSVKAEQYLLENE 60
|||||
Qy 61 TTKNYLGIDGIPBFGRCCTQELLFKGKSALINDKRAQTAPPGTGAIRVAADFLAKNTSV 120
|||||

Db 61 TTKNYLGIDGIPBFGRCCTQELLFGKGSALINDKRAARTQTPGCTGALRVAADFIAKNTSV 120
Qy 121 KRWWNSPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 121 KRWWNSPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGMPLPFDFAIYQGFARGLEDAEGLRAFAAMHKEIIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGMPLPFDFAIYQGFARGLEDAEGLRAFAAMHKEIIV 240
Qy 241 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVAATILSN 300
Db 241 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVAATILSN 300
Qy 301 DALRAIWEQELTDMRQRIORMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 360
Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 2
US-09-489-039A-12073
; Sequence 12073, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12073
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12073

Query Match 94.4%; Score 1931; DB 2; Length 402;
Best Local Similarity 93.4%; Pred. No. 3.6e-193;
Matches 370; Conservative 12; Mismatches 14; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETGTPTVLTSVKKAQVYLLNE 60
Db 7 MFENITAAPADPILGLADLFRADDRPEKINLGIVGYKDETGTPTVLTSVKKAQVYLLNE 66
Qy 61 TTKNYLGIDGIPBFGRCCTQELLFGKGSALINDKRAARTQTPGCTGALRVAADFIAKNTSV 120
Db 67 TTKNYLGIDGIPBFGRCCTQELLFGKGNAAIDKRAARTQTPGCTGALRVAADFIAKNTSV 126
Qy 121 KRWWNSPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 127 KRWWNSPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 186
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGMPLPFDFAIYQGFARGLEDAEGLRAFAAMHKEIIV 240
Db 187 CHNPTGIDPTLEQWOTLAQLSVEKGMPLPFDFAIYQGFARGLEDAEGLRAFAAMHKEIIV 246
Qy 241 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVAATILSN 300
Db 247 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVAATILSN 306
Qy 301 DALRAIWEQELTDMRQRIORMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 360
Db 307 DALRAIWEQELTDMRQRIORMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 366
Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

Db 367 LREEFAIYAVASGRINRVAGMTPDNMAPLCEAIVAVL 402
RESULT 3
US-09-252-991A-28999
; Sequence 28999, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28999
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28999
Query Match 48.0%; Score 982; DB 2; Length 413;
Best Local Similarity 48.0%; Pred. No. 7.9e-94;
Matches 190; Conservative 70; Mismatches 136; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVGYKDETGTPTVLTSVKKAQVYLLNE 60
Db 18 LFSAVENAPDRPILGLNEAFNADTRPGKINLGIVGYNEESGRIPLLRAVQAEEKARIEAH 77
Qy 61 TTKNYLGIDGIPBFGRCCTQELLFGKGSALINDKRAARTQTPGCTGALRVAADFIAKNTSV 120
Db 78 APRGYLPLEGIAAYDQGVQKLLFGNESELAAAGRVVTTQAVGGTGALKGADFLKRLLPD 137
Qy 121 KRWWNSPNHKSVPNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 138 ATVAISDPSENHRALEFAAGFPVQNYRYDDAASNGVNRAGLLEDLALPARSIVLVHAC 197
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGMPLPFDFAIYQGFARGLEDAEGLRAFAAMHKEIIV 240
Db 198 CHNPTGVDELDDWKQVLDVILKAKGHVFFLDIAVQGFNGEEDAAAARVLFQAQGLSPFV 257
Qy 241 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVAATILSN 300
Db 258 SSSFSKFSLSYGERVAGALSIVTESRDESARVLSQVKRVIRTNYSNPPHSGSVSSVLS 317
Qy 301 DALRAIWEQELTDMRQRIORMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 360
Db 318 PELRALWEQELGEMRDIRDMRLAMVQLAAHGAHAKRDFSFVGRQGMFSYGLTADQVER 377
Qy 361 LREEFGYVAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 378 LKTEFGIYAVSTGRIICVAALNKSLEITIKAIQVVL 413

RESULT 4
US-09-252-991A-29078
; Sequence 29078, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27


```

; SEQ ID NO 8472
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-8472

Query Match 41.3%; Score 844.5; DB 2; Length 445;
Best Local Similarity 42.1%; Pred. No. 2.2e-79;
Matches 167; Conservative 74; Mismatches 155; Indels 1; Gaps 1;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPVLTSVKKAEOYL-LEN 59
Db 49 VFQKVDAYAGDPILSLMERFERDEPRSDKVNLSIGLYYNDGIIIPOLQVAEAEARLNAEP 108

Qy 60 ETTKNYLGIDGIPFGRCCTQELLFGKGSALINDKRAARTQATPGGTGALRVAADFLAKNTS 119
Db 109 HGASLYLPMEGSLGVRQIAPELLFGAHTALKQNRIASIQTVGGSGALKVGADEFLKRYFP 168

Qy 120 VKRVVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLPHG 179
Db 169 ESHVWVSDPTWENHIAIFEGAGFEVSTVPWFDKATNGVRPFDLLATLQTLPARDIVLLHP 228

Qy 180 CCHNPTGIDPTLEQWQTLAQLSVKGLPLPFDPAYQGFARGLEDEAGLRAFAAMHKELI 239
Db 229 CCHNPTGADLTAPQWDRVVEVLKARQLIPFLDIAYQGGGGLKEDKAYAIRAIASAGMPML 288

Qy 240 VASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILS 299
Db 289 VNSFSKIFSLYGERVGLSVVCEDESETAGRVQLKATVRRNYSPPSFGAQVVAIVLN 348

Qy 300 NDALRAIWEQELTDMRQRIQRMRLQFVNTLQEKGNRDFSFIKQNGMFSFGLTKEQVL 359
Db 349 DAGLKATQAEVDAMRAHILTMRQALVDALQVAPGSKVDYLLKQRGMSYTGFSAAQVD 408

Qy 360 RLREFFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
Db 409 RLREDFGVYLIASGRMCVAGLNSRNVOQVAKAFVAVM 445

RESULT 8
US-09-543-681A-7295
; Sequence 7295, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7295
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-7295

Query Match 41.2%; Score 843.5; DB 2; Length 419;
Best Local Similarity 41.9%; Pred. No. 2.5e-79;
Matches 166; Conservative 75; Mismatches 154; Indels 1; Gaps 1;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPVLTSVKKAEOYL-LEN 59
Db 23 VFQVQAEAFAGDPILSLMDVYNNKDPQDKINLSIGLYYDEEGKTPILGTVSVAQQLNANWT 82

Qy 60 ETTKNYLGIDGIPFGRCCTQELLFGKGSALINDKRAARTQATPGGTGALRVAADFLAKNTS 119
Db 83 PTATLYLPMEGSLGVRQIAPELLFGAHTALKQNRIASIQTVGGSGALKVGADEFLKRYFP 142

Qy 120 VKRVVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLPHG 179

```

```

Db 143 SSEVWISDPTWHDNHSIFAGSGFKVNYYPFDPETKGVKFDALIDCFKKLPEKSIIVLMHP 202
Qy 180 CCHNPTGIDPTLEQWQTLAQLSVKGLPLPFDPAYQGFARGLEDEAGLRAFAAMHKELI 239
Db 203 CCHNPTGSDLTAKQMDQVTEILKARQAIPFLDIAYQGFABSLDDDDAYAVRAMAKAGLPVL 262
Qy 240 VASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILS 299
Db 263 VNSFSKIFGIYGERVGLSVVCDNAKECEHVJGQJKAARRIYSSPANYGAQILVQVLS 322
Qy 300 NDALRAIWEQELTDMRQRIQRMRLQFVNTLQEKGNRDFSFIKQNGMFSFGLTKEQVL 359
Db 323 DHVLTAKQWKEVAHMDRIKEMRVTLVNLKEALPEKPNFDHLLTQRMFSYTGFSPEQVD 382
Qy 360 RLREFFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 395
Db 383 RLREFFGVLIGTGRVCMAGVNNNNVQRIQAFAAV 418

RESULT 9
US-09-328-352-4895
; Sequence 4895, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4895
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-4895

Query Match 40.6%; Score 830.5; DB 2; Length 427;
Best Local Similarity 42.1%; Pred. No. 6e-78;
Matches 167; Conservative 72; Mismatches 157; Indels 1; Gaps 1;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPVLTSVKKAEOYL-LEN 59
Db 24 MFQHIPYPYAGDPILSLMEQFNADTRSEKVNLSIGLYNEDSIVPQLETTIEAQKRIEFPKN 83
Qy 60 ETTKNYLGIDGIPFGRCCTQELLFGKGSALINDKRAARTQATPGGTGALRVAADFLAKNTS 119
Db 84 GKTKLYLPMEGFKPYREAIQALLFGANS PAVKAGRAVTIQTLLGGSGALKVGADEFLKTYFP 143
Qy 120 VKRVVSNPSPNPKSVNSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLPHG 179
Db 144 NSDVVWSOPTWHDNHSIFAGSGFKVNYYPFDPETKGVKFDALIDCFKKLPEKSIIVLMHP 203
Qy 180 CCHNPTGIDPTLEQWQTLAQLSVKGLPLPFDPAYQGFARGLEDEAGLRAFAAMHKELI 239
Db 204 CCHNPTGADLNPAQMDQVIAVLKDRNLIPFLDIAYQGGGMEEDAYAIRALDQAGLNFI 263
Qy 240 VASSYSKNFGLYNERVGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILS 299
Db 264 VNSFSKIFSLYGERVGLTFVCDNAEAAQCTFQOLKATVRRIYSSPPPTTGAWLVDEVLN 323
Qy 300 NDALRAIWEQELTDMRQRIQRMRLQFVNTLQEKGNRDFSFIKQNGMFSFGLTKEQVL 359
Db 324 DAELNQOQWQGVKEMRERIINKRSILKDELTKALPDRDFSILVQNGMFSYTGLTAEQVD 383
Qy 360 RLREFFGVYAVASGRVNVAGMTDPNNAPLCEAIVAVL 396
Db 384 ILREYAIYLRSGRICVAGLNNNNVTVAKAMAEVL 420

RESULT 10
US-09-976-594-430
; Sequence 430, Application US/09976594

```


Qy 180 CCHNPTGIDPTLEQWQTLAQLSVEKGLPLDFDFAVQGFARGLEDEAGLRAFAAMHKELI 239
 Db 178 CCHNPTGVDFTEDMWEKELAILMKEKRELIPFDFAVQGFARGLEDEAGLRAFAAMHKELI 237
 Qy 240 VASSYKNGFLYNERVAGTVAADSETVDRAFSQMKAAIRANYSNPPAHGASVATILS 299
 Db 238 VAASSKKNFALYGERVGVFAVHSTFTDELVKIHSFLEEKIRGEYSSPQWGWVEIVSTILS 297
 Qy 300 NDALRAIWEQELTDMRQRIQMRQLFVNTLOEKGANRDFSFIIKQNGMFSFGLTKEQVL 359
 Db 298 NPYLKEEQWSELNFIRESLGKWRTRFVQALR-KVAGHTFDELLSQHGFAYPGFSDKQVL 356
 Qy 360 RLREFFGYVAVASGRVNVAGTMPDNMAPLCEAIV 393
 Db 357 FLREQHAVYTTAGGRMNLGITEKNIDHWQSF 390

RESULT 13

US-09-438-185A-742
 ; Sequence 742, Application US/09438185A
 ; Patent No. 6822071
 ; GENERAL INFORMATION:
 ; APPLICANT: Stephens, Richard
 ; APPLICANT: Mitchell, Wayne
 ; APPLICANT: Kalman, Sue
 ; APPLICANT: Davis, Ronald
 ; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
 ; FILE REFERENCE: 018941-000411US
 ; CURRENT APPLICATION NUMBER: US/09/438,185A
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: US 60/108,279
 ; PRIOR FILING DATE: 1998-11-12
 ; PRIOR APPLICATION NUMBER: US 60/128,606
 ; PRIOR FILING DATE: 1999-04-08
 ; NUMBER OF SEQ ID NOS: 1074
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 742
 ; LENGTH: 397
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; OTHER INFORMATION: CPn0740
 US-09-438-185A-742

Query Match 35.5%; Score 726.5; DB 2; Length 397;
 Best Local Similarity 39.6%; Pred. No. 4.1e-67; Indels 9; Gaps 3;
 Matches 156; Conservative 75; Mismatches 154;
 Qy 2 PENITAAPADPILGLADLPRADERPGKINLGIVYKDETGTPTVLTSVKKAQYILLENET 61
 Db 6 FNHIPTFSDALILGLQNVFFADKREKYNLVIGVYEHQKRYGGLSCIRKAQTVILESEQ 65
 Qy 62 TKNYLGIDGIPFGCTQELLFG--KGSALINDKRAQTAPGCTGALRVAADFLAKNTS 119
 Db 66 NKSYPISGLQIFLDMRELVEGAVDPSAIVG-----FQSLGGTCAHLGARLLSVAKG 119
 Qy 120 VKRVVNSPNSWNHKSVEFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHG 179
 Db 120 SGKVVPEQTWGNHRIFSQEGLEVIRPYYSKEQQLFEPFLAFLKEVEKNSVILLHG 179
 Qy 180 CCHNPTGIDPTLEQWQTLAQLSVEKGLPLDFDFAVQGFARGLEDEAGLRAFAAMHKELI 239
 Db 180 CCHNPTGVDFTEDMWEKELAILMKEKRELIPFDFAVQGFARGLEDEAGLRAFAAMHKELI 239
 Qy 240 VASSYKNGFLYNERVAGTVAADSETVDRAFSQMKAAIRANYSNPPAHGASVATILS 299
 Db 240 VAASSKKNFALYGERVGVFAVHSTFTDELVKIHSFLEEKIRGEYSSPQWGWVEIVSTILS 299
 Qy 300 NDALRAIWEQELTDMRQRIQMRQLFVNTLOEKGANRDFSFIIKQNGMFSFGLTKEQVL 359
 Db 300 NPYLKEEQWSELNFIRESLGKWRTRFVQALR-KVAGHTFDELLSQHGFAYPGFSDKQVL 358

Qy 360 RLREFFGYVAVASGRVNVAGTMPDNMAPLCEAIV 393
 Db 359 FLREQHAVYTTAGGRMNLGITEKNIDHWQSF 392

RESULT 14

US-09-801-874-5
 ; Sequence 5, Application US/09801874
 ; Patent No. 6582935
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN AMINOTRANSFERASE
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
 ; TITLE OF INVENTION: AMINOTRANSFERASE PROTEINS,
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: CLO00615
 ; CURRENT APPLICATION NUMBER: US/09/801,874
 ; CURRENT FILING DATE: 8001-03-09
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-801-874-5

Query Match 33.9%; Score 694; DB 2; Length 365;
 Best Local Similarity 43.2%; Pred. No. 8.9e-64;
 Matches 153; Conservative 47; Mismatches 144; Indels 10; Gaps 5;

Qy 10 ADPIL--GLADLPRADERPGKINLGIVYKDETGTPTVLTSVKKAQYIL--ENETTKNYL 66
 Db 9 AQPVLVFKLTADREFDPDKVNLGVGAYRTDCHPWLPVVKVQEKIANDNSLNHLYL 68
 Qy 67 GIDGIPFGCTQELLFGKGSALINDKRAQTAPGCTGALRVAADFLAK----NTSVK 121
 Db 69 PILGAEFRCASRLAGDDSPALKEKRVGVQSLGGTGLRIGADFLRWYNGTNNKNT 128
 Qy 122 RVVYNSFPNHSKVFNSAGL--EVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHG 180
 Db 129 PVYVSSPTWENHNAVFAAGFKDIRSYRWDAEKRLDQLQGLNDLENAPESFIVVLHAC 188
 Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGLPLDFDFAVQGFARG--LEEDAEGLRFAAMHKELI 239
 Db 189 AHNPTGIDPTPEQWKQIASVMKRFLPFPFFDSAYQGFASGNLERDAWAIKRVFVSGEFF 248
 Qy 240 VASSYKNGFLYNERVAGTVAADSETVDRAFSQMKAAIRANYSNPPAHGASVATILS 299
 Db 249 CAQSFSGKNGFLYNERVAGTVAADSETVDRAFSQMKAAIRANYSNPPAHGASVATILS 308
 Qy 300 NDALRAIWEQELTDMRQRIQMRQLFVNTLOEKGANRDFSFIIKQNGMFSFGL 353
 Db 309 NPFLFEETGTVNKTMDRILTMRSLELRALKALTPGTWNHITDQIGMFSFTGL 362

RESULT 15

US-09-248-796A-17477
 ; Sequence 17477, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 17477

US-09-248-796A-17477
 ; Sequence 17477, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 17477

; LENGTH: 303

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-17477

```

Query Match      26.5%; Score 542; DB 2; Length 303;
Best Local Similarity 43.5%; Pred. No. 5.3e-48;
Matches 114; Conservative 35; Mismatches 101; Indels 12; Gaps 3;

QY      2 FENITAAPADPILGLADLFRADERPGKINLIGIVYKDETGTPTVLTSVKKAQYLLNET 61
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
36 WNDIPLAPDKILGISEAYNNDSPQKINLGVGYARDNSGKPIIPPSVKKAEIILGKET 95
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
QY      62 TKNYLGIDGIPFGRCTOELLF-----GKSALINDKRARTAPGPGTGALRVAADFIA 115
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
96 EKEYTAIVGSKNFQIVKRFIENNSKNDANGKQLIDGGRIVTAQTISGTGSLRVIADFLN 155
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
QY      116 KNTSVKRVVWVSNPSPNHSVFNHSAGLEVREYAYDAENHTLDPDALINSLNEAQAGDVV 175
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
156 RFYSNKKILVPKPTWANHVAVFKDAGLEPEFYYSYETSKNDLDYANLKKSLTAAPEGSIV 215
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
QY      176 LFHGGCHNPTGIDPTLEQWOTLAQLSVEKGWLPDFEFAYQGFARGLE-EDAEGLRAFAAM 234
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
216 LLHACHNPTGMDLISEQWDEVLIVQDKKFPFLVDMAYQGPASGKPFEDIGLIRKLTKL 275
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
QY      235 HKE-----LIVASSYSKNFGLY 251
Db      : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : |||
276 ANENKIPSFALCQSFKNMGly 297

```

Search completed: March 14, 2006, 18:01:38
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 18:11:25 ; Search time 166 Seconds
(without alignments)
996.749 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITRAPADPILGLADLF.....VAGTMDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing, first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2045	100.0	396	4	US-10-369-493-785	Sequence 785, App
2	2045	100.0	396	4	US-10-673-786A-2	Sequence 2, Appli
3	1654	80.9	396	4	US-10-369-493-21125	Sequence 21125, A
4	1318	64.4	322	4	US-10-369-493-231	Sequence 231, App
5	1311	64.1	397	4	US-10-275-036A-188	Sequence 188, App
6	1005.5	48.2	394	4	US-10-369-493-11824	Sequence 11824, A
7	996	48.7	396	4	US-10-369-493-7497	Sequence 7497, Ap
8	996	48.7	398	4	US-10-369-493-4738	Sequence 4738, Ap
9	982	48.0	397	4	US-10-369-493-8424	Sequence 8424, Ap
10	975	47.7	397	4	US-10-369-493-15828	Sequence 15828, A
11	975	47.7	397	4	US-10-369-493-16208	Sequence 16208, A
12	975	47.7	398	4	US-10-369-493-15460	Sequence 15460, A
13	947	46.3	398	4	US-10-369-493-14025	Sequence 14025, A
14	924	45.2	395	4	US-10-369-493-9458	Sequence 9458, Ap
15	917	44.8	395	4	US-10-369-493-9222	Sequence 9222, Ap
16	917	44.8	420	4	US-10-425-114-63613	Sequence 63613, A
17	914	44.7	420	4	US-10-425-114-63242	Sequence 63242, A
18	914	44.7	459	4	US-10-767-701-46300	Sequence 46300, A
19	913	44.6	423	4	US-10-425-114-57028	Sequence 57028, A
20	911.5	44.6	395	4	US-10-369-493-13723	Sequence 13723, A
21	911	44.5	417	4	US-10-425-114-68709	Sequence 68709, A
22	909	44.4	400	4	US-10-369-493-17487	Sequence 17487, A
23	904.5	44.2	464	4	US-10-425-115-197293	Sequence 197293, A
24	893	43.7	396	4	US-10-369-493-8659	Sequence 8659, Ap
25	885	43.3	456	4	US-10-424-599-239081	Sequence 239081, A
26	884.5	43.3	396	4	US-10-369-493-21154	Sequence 21154, A
27	875.5	42.8	396	4	US-10-369-493-287	Sequence 287, App

28	870	42.5	1991	4	US-10-437-963-187401	Sequence 187401, A
29	866	42.3	412	4	US-10-425-115-287905	Sequence 287905, A
30	866	42.3	574	4	US-10-425-115-363667	Sequence 363667, A
31	865	42.3	520	4	US-10-425-115-363669	Sequence 363669, A
32	863	42.2	463	4	US-10-424-599-172508	Sequence 172508, A
33	861	42.1	459	4	US-10-767-701-46309	Sequence 46309, A
34	851	41.6	396	4	US-10-369-493-4827	Sequence 4827, Ap
35	851	41.6	396	4	US-10-369-493-7586	Sequence 7586, Ap
36	846	41.4	413	5	US-10-481-032A-290	Sequence 290, App
37	843.5	41.2	397	4	US-10-369-493-23657	Sequence 23657, A
38	831.5	40.7	397	4	US-10-422-366-32	Sequence 32, Appl
39	831.5	40.7	397	5	US-10-925-216-32	Sequence 32, Appl
40	831.5	40.7	397	5	US-10-903-582-32	Sequence 32, Appl
41	831.5	40.7	397	5	US-10-918-401A-32	Sequence 32, Appl
42	831.5	40.7	397	5	US-10-969-245-32	Sequence 32, Appl
43	828	40.5	424	6	US-11-097-143-5940	Sequence 5940, Ap
44	825	40.3	403	5	US-10-739-930-6191	Sequence 6191, Ap
45	823	40.2	710	5	US-10-450-763-54363	Sequence 54363, A

RESULT 1
US-10-369-493-785
; Sequence 785, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 785
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-785

Query Match	100.0%	Score	2045	DB	4	Length	396
Best Local Similarity	100.0%	Pred. No.	4.2e-190				
Matches	396	Conservative	0	Mismatches	0	Indels	0
Gaps	0						

Qy	1	MFENITRAPADPILGLADLF	FRADRP	PGKIN	IGVYKDE	TGKT	PVLT	SVKKA	EOYLL	ENE	60
Db	1	MFENITRAPADPILGLADLF	FRADRP	PGKIN	IGVYKDE	TGKT	PVLT	SVKKA	EOYLL	ENE	60
Qy	61	TTKYL	GIDG	IP	FGRC	CTQ	ELL	FGK	GSAL	IND	KART
Db	61	TTKYL	GIDG	IP	FGRC	CTQ	ELL	FGK	GSAL	IND	KART
Qy	121	KRVV	SWSP	NHKS	VFN	SAG	LEV	REY	AYDA	ENHT	LD
Db	121	KRVV	SWSP	NHKS	VFN	SAG	LEV	REY	AYDA	ENHT	LD
Qy	181	CHN	PTG	ID	PT	LEQ	WT	LA	QLS	VEK	GW
Db	181	CHN	PTG	ID	PT	LEQ	WT	LA	QLS	VEK	GW
Qy	241	ASS	YSK	KN	FG	LN	ER	V	AG	CT	LV
Db	241	ASS	YSK	KN	FG	LN	ER	V	AG	CT	LV
Qy	301	DAL	RAI	WE	Q	EL	T	M	R	Q	R
Db	301	DAL	RAI	WE	Q	EL	T	M	R	Q	R


```
/ Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7497
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7497

Query Match          48.7%; Score 996; DB 4; Length 396;
Best Local Similarity 48.2%; Pred No. 6.8e-88;
Matches 191; Conservative 67; Mismatches 138; Indels      0; Gaps    0;

Qy   1 MFENITAAPADPILGLADIPLFRADERPGKINLGIGYVVDKETGTPVLTSVKKAEGVLLNE 60
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        1 LFSAVELAPRDPILGLEAFNADRTRTKVNLVGVVFVEEKGIPLLRAVRDAEKARVDAA 60

Qy   61 TTKNYLGIQIPEFGRECTQLLFSGKSALINDKRARTAQTTGGTCALARVAADFIAKTSTV 120
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        61 LPRGYLPIEGIAAYDAAVKLLLGNDSPLJIAAGRVVTQAALGGTGALKIGADFLKLRLNP 120

Qy   121 KRYVVNSPNWPHSKSVFNESAGLEVREYAYYDAENTHLPFDALINSLEAQADVVLPHGC 180
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        121 AKVAISDPSENHRALFEGAGEVVSYPDYDHGTGVNFDMLSALSINYAAGTIVVLHAC 180

Qy   181 CHNPTGDIDPLEOWTLAQSVEKWGLPLDFPAYOGFARGLEEDEAEGLRAFAMHKELIV 240
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        181 CHNPTGDLNIEQMKVVEVKARNLVPPLDIAYOGFGDNTEADAATAVRLFAAAELNVFV 240

Qy   241 ASSYSKNFYNGERVACTIVAADSETVDRATFSOMKAIRANYSNPPAHGAASVATILSN 300
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        241 SSDFSXSFSLYGSRVGALSIIITASKEEARVLSQLKRVIRTNTSYSPPTHGGSVAAVLAS 300

Qy   301 DALRAETWEOLTDMRORIORMQLFVNTLOEKANRDSEFIIKONGMFSGLTKEOVL R 360
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        301 PELRATWETAEMRRRIIRAMRNGLVERUKASGVDRDFSVNAQRGMFSYGLTAPQVDR 360

Qy   361 LRREFGYVASGRVNAVGMTPDNMPCLEAI VAVL 396
Db       :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
        361 LRREFGYIYSTGRCVCAALTNLNDVANATHVL 396


RESULT 8
US-10-369-493-4738
; Sequence 4738, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4738
```


[illegible]

Job time : 168 secs

```
Db 181 CHNPTGADLSRDQWQKQVALFOERQLLPCIDLAYQGFNOGIDADAYAIRLLAEGISNVV 240
Qy 240 VASSYSKNFGLYNERVACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
Db 241 VANSYSKFSGLYGERVGLSIVASNTEQAQAIQSQVKRIIRTIYSSPPAHGAYLVAGVLN 300
Qy 300 NDALRAIWEQELTDMRQIRQRLFVNTLQEKGNRDPFSFIKONGMFSFSGLTKEQVL 359
Db 301 SHELRTLWEQELTQMRRIHGLRAGLVRLKALGA-PEPDFIQRQAGMFSYSGLSKIQVD 359
Qy 360 RLREBFYGVAVASGRVNVAGTTPDNMAPLCEAIVAV 395
Db 360 RLREBFYGVAVSSGRICVAALSQHKLEYVAQAVKV 395

RESULT 15
US-10-369-493-9222
; Sequence 9222, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9222
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9222

Query Match 44.8%; Score 917; DB 4; Length 395;
Best Local Similarity 45.5%; Pred. No. 3.4e-80;
Matches 180; Conservative 79; Mismatches 135; Indels 2; Gaps 2;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGVYKDETGTGTPVLTSVKKAEOYLLENE 60
Db 1 LFTDVELVPGPILSLNDAYNADTTNKVNLGIGYCDSEGCIPLLRAVQQVEEQAKHP 60
Qy 61 TTKNYLGIDGIPFGRCRTOELLFGKGSALINDKRARTATQPGTGALRVAADFLAKNTSV 120
Db 61 KPRGYLPIDGLPAYIKATQQLLFQVDSLLLTAGRVATSTQVGGSGALRVAAEVLKQVLPH 120
Qy 121 KRVVSNPSNHNKSVNSAGLEVREYAYDAENHTLDPDALINSNEAQAGDVVLFHGC 180
Db 121 ATVAISRPNSWNRHALFTAAGFKIEDTYFDTLGHAVDFTGMWADLAKLPKTVVLLHGC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWMLPLFPAYOGFARGLEEDAEGLRFAAAMH-KELI 239
Db 181 CHNPTGADLSRDQWQKQVALFOERQLLPCIDLAYQGFNOGIDADAYAIRLLAEGISNVV 240
Qy 240 VASSYSKNFGLYNERVACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILS 299
Db 241 VANSYSKFSGLYGERVGLSIVASNTEQAQAIQSQVKRIIRTIYSSPPAHGAYLVASVLN 300
Qy 300 NDALRAIWEQELTDMRQIRQRLFVNTLQEKGNRDPFSFIKONGMFSFSGLTKEQVL 359
Db 301 SHELRTLWEQELTQMRRIHGLRAGLVRLKALGV-PEPDFIQRQAGMFSYSGLSKIQVD 359
Qy 360 RLREBFYGVAVASGRVNVAGTTPDNMAPLCEAIVAV 395
Db 360 RLREBFYGVAVSSGRICVAALSQHKLEYVAQAVKV 395
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 18:12:15 ; Search time 22 Seconds
(without alignments)
501.028 Million cell updates/sec

Title: US-10-673-786A-2
Perfect score: 2045
Sequence: 1 MFENITAAPADPILGLADLF.....VAGMTDPNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications_AA New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2011	98.3	396	7	US-11-114-922-76 Sequence 76, Appl
2	1317	64.4	423	6	US-10-467-657-6536 Sequence 6536, Ap
3	842	41.2	397	6	US-10-467-657-1060 Sequence 1060, Ap
4	831.5	40.7	397	6	US-10-979-821-32 Sequence 32, Appl
5	831.5	40.7	397	6	US-11-114-922-32 Sequence 32, Appl
6	818	40.0	419	6	US-10-979-821-6 Sequence 6, Appli
7	818	40.0	419	6	US-11-114-922-6 Sequence 6, Appli
8	815	39.9	389	6	US-10-979-821-2 Sequence 2, Appli
9	815	39.9	389	7	US-11-114-922-2 Sequence 2, Appli
10	810	39.6	419	6	US-10-979-821-4 Sequence 4, Appli
11	810	39.6	419	7	US-11-114-922-4 Sequence 4, Appli
12	794	38.8	430	6	US-10-878-556A-146 Sequence 146, App
13	756	37.0	412	6	US-10-878-556A-149 Sequence 149, App
14	699	34.2	447	7	US-11-245-400-17 Sequence 17, Appli
15	632.5	30.9	412	6	US-10-979-821-8 Sequence 8, Appli
16	632.5	30.9	412	7	US-11-114-922-8 Sequence 8, Appli
17	153.5	7.5	429	6	US-10-793-626-3174 Sequence 3174, Ap
18	123.5	6.0	369	7	US-11-055-822-116 Sequence 116, App
19	123.5	6.0	369	7	US-11-055-822-486 Sequence 486, App
20	113	5.5	473	7	US-11-087-099-10299 Sequence 10299, A
21	111	5.4	452	6	US-10-467-657-1850 Sequence 1850, Ap
22	105.5	5.2	392	7	US-11-087-099-941 Sequence 941, App
23	105.5	5.2	397	7	US-11-232-405A-42 Sequence 42, Appl
24	102	5.0	473	7	US-11-087-099-1989 Sequence 1989, Ap
25	100.5	4.9	871	7	US-11-087-099-7516 Sequence 7516, Ap

26	99	4.8	393	6	US-10-979-821-10	Sequence 10, Appl
27	99	4.8	393	7	US-11-114-922-10	Sequence 10, Appl
28	98	4.8	399	7	US-11-096-568A-22461	Sequence 22461, A
29	98	4.8	432	7	US-11-096-568A-22460	Sequence 22460, A
30	96	4.7	434	7	US-11-096-568A-20937	Sequence 20937, A
31	96	4.7	462	7	US-11-096-568A-20936	Sequence 20936, A
32	96	4.7	505	7	US-11-096-568A-20935	Sequence 20935, A
33	96	4.7	623	7	US-11-072-512-2547	Sequence 2547, Ap
34	96	4.7	657	6	US-10-873-528-57	Sequence 57, Appl
35	95	4.6	389	6	US-10-467-657-4468	Sequence 4468, Ap
36	95	4.6	461	7	US-11-096-568A-10640	Sequence 10640, A
37	94.5	4.6	444	7	US-11-205-109-7	Sequence 7, Appli
38	94	4.6	395	7	US-11-087-099-8148	Sequence 8148, Ap
39	94	4.6	894	6	US-10-485-517-416	Sequence 416, App
40	93	4.5	395	7	US-11-087-099-11586	Sequence 11586, A
41	90	4.4	326	7	US-11-096-568A-22462	Sequence 22462, A
42	90	4.4	356	7	US-11-087-099-1503	Sequence 1503, Ap
43	89.5	4.4	394	6	US-10-510-386-208	Sequence 208, App
44	89.5	4.4	633	7	US-11-087-099-10840	Sequence 10840, A
45	88.5	4.3	395	6	US-10-467-657-3516	Sequence 3516, Ap

ALIGNMENTS

RESULT 1

US-11-114-922-76
; Sequence 76, Application US/11114922
; Publication No. US20050282260A1
; GENERAL INFORMATION:
; APPLICANT: HICKS, PAULA M.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0396
; CURRENT APPLICATION NUMBER: US/11/114,922
; CURRENT FILING DATE: 2005-04-26
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 76
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-114-922-76

Query Match 98.3%; Score 2011; DB 7; Length 396;
Best Local Similarity 98.5%; Pred. No. 1.1e-161;
Matches 390; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy	1	MFENITAAPADPILGLADLFRADPRPGKINIGIVYKDETKTPTVLTSVKKAQYLLNE	60
Db	1	MFENITAAPADPILGLADLFRADPRPGKINIGLYDETKIKIPVLTSVKKAQYLLNE	60
Qy	61	TKKYLIGDITPEGRCTQELLFGKGSALINDKARTAOPTGGTCALEVAADFLAKNTSV	120
Db	61	TKKYLIGDITPEGRCTQELLFGKGSALINDKARTAOPTGGTCALEVAADFLAKNTSV	120
Qy	121	KRVWVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC	180
Db	121	KRVWVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC	180
Qy	181	CHNPTGIDPTLEQWOTLAQLSVEKGLPLFDFAQGFARGLEEDAEGLRAFAAMHKEIV	240
Db	181	CHNPTGIDPTLEQWOTLAQLSVEKGLPLFDFAQGFARGLEEDAEGLRAFAAMHKEIV	240
Qy	241	ASSYSKNFGLYNERVGACTIVAAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN	300
Db	241	ASSYSKNFGLYNERVGACTIVAAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN	300

```

Qy 301 DALRAIWEQELTDMRQRIORMQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 360

Qy 361 LREFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 2
US-10-467-6536
; Sequence 6536, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6536
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-6536
Query Match 64.4%; Score 1317; DB 6; Length 423;
Best Local Similarity 61.9%; Pred. No. 3.3e-103; Indels 0; Gaps 0;
Matches 244; Conservative 62; Mismatches 88;
Qy 2 FENITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAQVLLNET 61
Db 30 FKHIEAAPADPILGLGEAFKAETPEKVNGLIGVYKDSAGATPIVKAKEAEKLLSEET 89
Qy 62 TKNYLIGIDIPFGRCCTQELLFGKGSALINDKRAKTAQTPGCTGALRVAADFLAKNTSVK 121
Db 90 TKNYLTIDGVADYNEQTOILLFGKDHEIYASRRAKTAQSLGCTGALRIAABFAKRLNAQ 149
Qy 122 RVVSNPSWPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGCC 181
Db 150 TIWISNTPWPNHNAIKAVGIQDKFYRYIDAKHCLDWDGHWIEDLNQAKQGDIVLLHGCC 209
Qy 182 HNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRAFAAMHKLIVA 241
Db 210 HNPTGIDPTPEQWETLAKLSAEKGWLPFLDFAYQGFNGLEEDAYGLRVFLKHNTLLIA 269
Qy 242 SSYSKNFGLYNERVGACTIONVAADSETVDRAFQSKAAAIRANYSNPPAHGASVVAATILSND 301
Db 270 SSYSKNFGWYNERVGAFTVAEDDEETAAARASHQIKTIIRTLYSNPASHGANTIALVLKND 329
Qy 302 ALRAIWEQELTDMRQRIORMQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 361
Db 330 DUKAQWIAELDEMRGRIKAMQKQVLELLKAKGSTQDFDPIIEQNGMFSFSLTPEQVRL 389
Qy 362 REEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAV 395
Db 390 KNEFAIYAVSRGINVAGITDDNIDYLCESIVK 423

```

```

RESULT 3
US-10-467-657-1060
; Sequence 1060, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita

```

```

; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1060
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1060
Query Match 41.2%; Score 842; DB 6; Length 397;
Best Local Similarity 41.7%; Pred. No. 2.9e-63;
Matches 165; Conservative 68; Mismatches 163; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETGTPTVLTSVKKAQVLLNET 60
Db 1 MYRHIEVYVPGDPILSLVETEFKNDPRPEKVNLSIGIYFDEGRMPVLESVSRAETARAAP 60
Qy 61 TKNYLIGIDIPFGRCCTQELLFGKGSALINDKRAKTAQTPGCTGALRVAADFLAKNTSV 120
Db 61 APSPYLPMGLDVYRSVAVQHLFLFGKGNPALAQQGRIVTVQTLGGSGALKVGADFLLHRWPE 120
Qy 121 KRVVSNPSWPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSNEAQAGDVVLFHGC 180
Db 121 ARAYVSQPTWDNHRGIFEGAGFEVGYTPYDPAIVGVKFDGMTAFFNTLPENSVLILHPC 180
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYQGFARGLEEDAEGLRAFAAMHKLIV 240
Db 181 CHNPTGVDMSERQWDEVLIQIKTRKLIPFMDIAYQGFGLDSDAYAVRKAVERMDLPLFV 240
Qy 241 ASSYSKNFGLYNERVGACTIONVAADSETVDRAFQSKAAAIRANYSNPPAHGASVVAATILSN 300
Db 241 SNSFSKNLSYGERVGGVSVCPNKEADLVFGQLKFTVRIYSSPPAHGAYIAADVNS 300
Qy 301 DALRAIWEQELTDMRQRIORMQLFVNTLQKGNRDFSFIIKQNGMFSFSLTKQVLR 360
Db 301 SELVALQWNEVYMRDRIRAMRQKLYGVLTARIPDRDFTFYFIKQNGMFGYGLSVGVRR 360
Qy 361 LREFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LRDEFAYVLLDSGRMVCVAGLNTSNITVYVADALAEVL 396

RESULT 4
US-10-979-821-32
; Sequence 32, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77

```

Db	61	HGASLYLPM	EGNLCYRHA	IAPI	LLFGADHP	LVKQOR	VATIT	QTITGG	SALKV	GADFLK	RYP	120				
Qy	120	VKRVVNSP	SWPNHKS	VNSAGL	EVREYAY	YDAENH	TLDF	ALIN	SLNEA	QAGD	VVLPHG	179				
Db	121	ESGVVSDP	TWENRVA	IFAGAG	FEVSTY	PWYDEAT	NGVRFN	DLAT	LTPL	PARSIV	LLHP	180				
Qy	180	CCHNPTGI	DPTELEQ	WOTLAQ	LSVEKG	WLPFL	PDFA	YQGF	ARGLE	EDAE	GLRPA	FAAMHKE	239			
Db	181	CCHNPTGA	DULTND	QWDV	IEILKAR	ELIPFL	DIAYQ	GFGAG	MEEDAY	AI	RAIASA	GLPAL	240			
Qy	240	VASSYKFN	GLYNER	VGACTL	VAA	DS	ETVDR	AFSOM	KAAI	RAIRAN	YSPN	PPAHGASV	VATILS	299		
Db	241	VSNSFSKI	ESLGER	VGGLSV	NCEDAE	AAAG	RVLG	LQKAT	VRNYS	SPN	PPFGA	VQVVA	AVLN	300		
Qy	300	NDALRAIW	EQELTDM	RQIR	QRLF	VNTL	QK	GANR	DFS	FTIK	ONGM	FSFGLT	KEQVL	359		
Db	301	DEALKASW	LA	VEEMR	TRILAM	RQELV	KVLS	TE	MPERN	FDYLLN	ORGM	FSYITGL	SAAQVD	360		
Qy	360	RLREEGVY	AVASGR	VNVAGM	TPDN	MAPL	CEA	IV	AVL					396		
Db	361	RLREEGVY	LIASGR	MCVAGL	TANV	QORV	AKA	FAA	AVM					397		
RESULT 6																
US-10-979-821-6																
; Sequence 6, Application US/10979821																
; Publication No. US20050244937A1																
; GENERAL INFORMATION:																
; APPLICANT: ABRAHAM, TIMOTHY W.																
; APPLICANT: CAMERON, DOUGLAS C.																
; APPLICANT: HICKS, PAULA M.																
; APPLICANT: MCFARLAN, SARA C.																
; APPLICANT: MILLIS, JIM																
; APPLICANT: ROSAZZA, JACK																
; APPLICANT: ZHAO, LISHAN																
; APPLICANT: WEINER, DAVID P.																
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE																
; FILE REFERENCE: 023829-0390																
; CURRENT APPLICATION NUMBER: US/10/979,821																
; CURRENT FILING DATE: 2004-11-03																
; PRIOR APPLICATION NUMBER: 10/422,366																
; PRIOR FILING DATE: 2003-04-23																
; PRIOR APPLICATION NUMBER: 60/374,831																
; PRIOR FILING DATE: 2002-04-23																
; NUMBER OF SEQ ID NOS: 77																
; SOFTWARE: PatentIn Ver. 3.3																
; SEQ ID NO 6																
; LENGTH: 419																
; TYPE: PRT																
; ORGANISM: Rhodobacter sphaeroides																
US-10-979-821-6																
Query Match 40.0%; Score 818; DB 6; Length 419;																
Best Local Similarity 42.9%; Pred. No. 3.3e-61;																
Matches 171; Conservative 61; Mismatches 157; Indels 10; Gaps 5																
Qy	1	MFENITAA	PADPIL	GLADL	FRADER	PGKIN	LIGI	GVYK	DET	GKT	PVLTS	VKKA	EQYL	LEN	60	
Db	26	MLTALKPQ	PADKIL	QIQF	REDAR	ADKID	LGV	GVYK	DP	TGLT	TPV	MRVA	AKAE	KELWE	85	
Qy	61	TTKNVIGI	DGIP	EGRC	CTOELL	FGKS	ALIN	D	KART	QA	TGCG	TGAL	RVA	ADFLAK	NTSV	120
Db	86	TTKTYTGL	AGEPAY	NAAMAK	LIL---	AGAV	PADR	VASV	AT	PGT	GA	VQA	QAL	ELIRMA	SPE	142
Qy	121	KRVVNSP	SWPNHKS	VNSAGL	EVREYAY	YDAENH	TLDF	ALIN	SLNEA	QAGD	VVLPHG					180
Db	143	ATVWISNP	TWPNHLS	IVKVLG	IPME	RYFY	DAET	GA	VD	AEGL	MED	LAQV	KAGD	VVLHGC		202
Qy	181	CCHNPTGI	DPTELEQ	WOTLAQ	LSVEKG	WLPFL	PDFA	YQGF	ARGLE	EDAE	GLRPA	FAAMHKE	LIV			240
Db	203	CCHNPTGA	NPNVOM	LVAVCES	LART	GA	VLID	LA	YQGF	D	GLE	MD	AAA	TRLLAT	RLPEVL	262
US-11-114-922-32																
; Software: PatentIn Ver. 3.3																
; SEQ ID NO 32																
; LENGTH: 397																
; TYPE: PRT																
; ORGANISM: Escherichia coli																
US-10-979-821-32																
Query Match 40.7%; Score 831.5; DB 6; Length 397;																
Best Local Similarity 42.1%; Pred. No. 2.3e-62;																
Matches 167; Conservative 74; Mismatches 155; Indels 1; Gaps 1;																
Qy	1	MFENITAA	PADPIL	GLADL	FRADER	PGKIN	LIGI	GVYK	DET	GKT	PVLTS	VKKA	EQYL	LEN	59	
Db	1	VFOKVD	YAGDP	PILTLM	ERFKED	PRSDK	VNLSIG	LYN	EDGII	POLQ	AVAE	AEARL	NAQP		60	
Qy	60	ETTKNVIGI	DGIP	EGRC	CTOELL	FGKS	ALIN	D	KART	QA	TGCG	TGAL	RVA	ADFLAK	NTS	119
Db	61	HGASLYLPM	EGNLCYRHA	IAPI	LLFGADHP	LVKQOR	VATIT	QTITGG	SALKV	GADFLK	RYP					120
Qy	120	VKRVVNSP	SWPNHKS	VNSAGL	EVREYAY	YDAENH	TLDF	ALIN	SLNEA	QAGD	VVLPHG					179
Db	121															

Qy 241 ASSYSKNFLYNERVGACTLV--AADSETVDRAFQSMKAAIRANYSNPPAHGASVATIL 298
Db 263 AASCCKNFIYRERTGILIAIEAGAGRTVQ---ANLFLNRQNYSPFPDGHGRLVTMIL 319
Qy 299 SNDALRAIWEQBLTDMRQIORMRQLFVNTLQ-EKGANRDFSFIKQNGMFSFSGLTKEQ 357
Db 320 EDETLSADWKAELEEVRLNMLTLRQLADALQAETGSRN-RGFVAEHRGMFSRLGITPAE 378
Qy 358 VLRLEEEFGVYAVASGRVNVAGMTDPNNMPLCEAIVAVL 396
Db 379 VERLRTHEGVYMGVDSRLNIAGLNRTTVPVLARAVAKVL 417

RESULT 7

US-11-114-922-6

; Sequence 6, Application US/11114922

; Publication No. US20050282260A1

; GENERAL INFORMATION:

; APPLICANT: HICKS, PAULA M.

; APPLICANT: MCFARLAN, SARA C.

; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE

; PRODUCTION OF MONATIN AND ITS PRECURSORS

; FILE REFERENCE: 023829-0396

; CURRENT APPLICATION NUMBER: US/11/114,922

; PRIOR FILING DATE: 2005-04-26

; PRIOR FILING DATE: 2003-04-23

; PRIOR FILING DATE: 2003-04-23

; PRIOR FILING DATE: 2002-04-23

; NUMBER OF SEQ ID NOS: 91

; SOFTWARE: Patent in Ver. 3.3

; SEQ ID NO 6

; LENGTH: 419

; TYPE: PRT

; ORGANISM: Rhodobacter sphaeroides

US-11-114-922-6

Query Match 40.0%; Score 818; DB 7; Length 419;

Best Local Similarity 42.9%; Pred. No. 3.3e-61;

Matches 171; Conservative 61; Mismatches 157; Indels 10; Gaps 5;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETGTPTLVTSVKKAEQVLLNE 60
Db 26 MLTALKPQPADKILQIQMFREDARADKIDILGVGVYKDPGLTPVMRAVKAERKLEWE 85
Qy 61 TTKNYLGIDGIPFGCRCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 86 TTKTYTGLAGEPAYNANAKLIL---AGAVPADRVASVATPGTGAVRQALELRMASPE 142
Qy 121 KRVMVSNPSPNHNKSVNSAGLEVREYAYYDAENHTLDFDALINSLNBAQAGDVVLFHGC 180
Db 143 ATVWISNPTWPNHLSIVKYLIGIPMEYRYFDAETCAVDAEGLMEDLAQVKAGDVVLLHGC 202
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGMPLDFPAYOGFARGLEEDAEGLRAFAAMHKLIV 240
Db 203 CHNPTGANPNVQWLAVCESLARTGAVPLIDLAYQGGDGLMEDMAAATRLATRLPEVLI 262

Qy 241 ASSYSKNFLYNERVGACTLV--AADSETVDRAFQSMKAAIRANYSNPPAHGASVATIL 298
Db 263 AASCCKNFIYRERTGILIAIEAGAGRTVQ---ANLFLNRQNYSPFPDGHGRLVTMIL 319
Qy 299 SNDALRAIWEQBLTDMRQIORMRQLFVNTLQ-EKGANRDFSFIKQNGMFSFSGLTKEQ 357
Db 320 EDETLSADWKAELEEVRLNMLTLRQLADALQAETGSRN-RGFVAEHRGMFSRLGITPAE 378
Qy 358 VLRLEEEFGVYAVASGRVNVAGMTDPNNMPLCEAIVAVL 396
Db 379 VERLRTHEGVYMGVDSRLNIAGLNRTTVPVLARAVAKVL 417

RESULT 8

US-10-979-821-2

; Sequence 2, Application US/10979821

; Publication No. US20050244937A1

; GENERAL INFORMATION:

; APPLICANT: ABRAHAM, TIMOTHY W.

; APPLICANT: CAMERON, DOUGLAS C.

; APPLICANT: HICKS, PAULA M.

; APPLICANT: MCFARLAN, SARA C.

; APPLICANT: MILLIS, JIM

; APPLICANT: ROSAZZA, JACK

; APPLICANT: ZHAO, LISHAN

; APPLICANT: WEINER, DAVID P.

; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE

; PRODUCTION OF MONATIN AND ITS PRECURSORS

; FILE REFERENCE: 023829-0390

; CURRENT APPLICATION NUMBER: US/10/979,821

; PRIOR FILING DATE: 2004-11-03

; PRIOR FILING DATE: 2003-04-23

; PRIOR FILING DATE: 2003-04-23

; PRIOR FILING DATE: 2002-04-23

; NUMBER OF SEQ ID NOS: 77

; SOFTWARE: Patent in Ver. 3.3

; SEQ ID NO 2

; LENGTH: 389

; TYPE: PRT

; ORGANISM: Sinorhizobium meliloti

US-10-979-821-2

Query Match 39.9%; Score 815; DB 6; Length 389;

Best Local Similarity 44.3%; Pred. No. 5.4e-61;

Matches 170; Conservative 63; Mismatches 137; Indels 14; Gaps 4;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINIGIVYKDETGTPTLVTSVKKAEQVLLNE 60
Db 1 MFDALARQADDPDLLFLGLFKDPERPGKVDLGVGVYRDETGTPTIFRAVKAERKLETTQ 60
Qy 61 TTKNYLGIDGIPFGCRCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 DSKAYTGPEDGLVPLDRLWELV---GGDTIERSHVAGVQTFGGSGALRLAADLIAR-MGG 116
Qy 121 KRVMVSNPSPNHNKSVNSAGLEVREYAYYDAENHTLDFDALINSLNBAQAGDVVLFHGC 180
Db 117 RGIWGLPSPNHNAPFKAAGLDIATYDFDIPSQSVIFONLVSALEGAASGDVALLHAS 176
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGMPLDFPAYOGFARGLEEDAEGLRAFAAMHKLIV 240
Db 177 CHNPTGVLSEAQWMEIATLVAERGLPLVDLAYQGGFGRGLDQDVAGLRHLLGVVPEALV 236
Qy 241 ASSYSKNFLYNERVGACTLVAAADSETVDRAFQSMKAAIRANYSNPPAHGASVATILSN 300
Db 237 AVSCSKSFGLYRERAGAI PARTSTASADRVRSNLAGARTSYSMPPDHGAUVRTILDD 296
Qy 301 DALRAIWEQBLTDMRQIORMRQLFVNTL---QEKGANRDFSFIKQNGMFSFSGLTKE 356
Db 297 PELARDWTEELTMRMTGLRRSLAEGRLTRWQSLGAVAD-----QEGMFSMLPLSEA 350
Qy 357 QVLRLEEEFGVYAVASGRVNVAGM 380
Db 351 EVMLRTEHGIYMPASGRINIAGL 374

RESULT 9

US-11-114-922-2

; Sequence 2, Application US/11114922

; Publication No. US20050282260A1

; GENERAL INFORMATION:

; APPLICANT: HICKS, PAULA M.

; APPLICANT: MCFARLAN, SARA C.

; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE

; PRODUCTION OF MONATIN AND ITS PRECURSORS

; FILE REFERENCE: 023829-0396

; CURRENT APPLICATION NUMBER: US/11/114,922

; CURRENT FILING DATE: 2005-04-26

Db 143 ATWISNTWPNHLSIVKVLGTPMREYRYFDAETGAVDAEGMEDLAQVKAGDVVLLHGC 202
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGLRPAFAAMHKEIV 240
Db 203 CHNPTGANPNVQWLAIACESLARTGAVPLIDLAVQFGDGLMDAAATRLLATRLPEVLI 262
Qy 241 ASSYSKNFGLYNERVGACTLV--AADSETVDRAFQSKAAIRANYSNPPAHGASVATIL 298
Db 263 AASCSKNFGIYRRTGIIAIGEAAGRGTQV---ANLNFLLARQNYVFPFDHGARLVNTWL 319
Qy 299 SNDALRAIWEQBLTDMRORIORMROLFVNTLQ-EKGANRDFSFIKQNGMFSFSGLTKEQ 357
Db 320 EDETUSADWKAELEVRNLMLTLRQLADALQAEATGSNR-PGFVAEHRGMSRSLGIIPAE 378
Qy 358 VLRLEEFQYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
Db 379 VERLRTERGVMVGVDSRLNIAGLNRTTVPVLARAVAKVL 417

RESULT 12

US-10-878-556A-146
; Sequence 146, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 146
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/aatm_human
; DATABASE ENTRY DATE: 1986-07-21
US-10-878-556A-146

Query Match 38.8%; Score 794; DB 6; Length 430;

Best Local Similarity 41.2%; Pred No. 3.6e-59;
Matches 161; Conservative 67; Mismatches 161; Indels 2; Gaps 2;

Qy 4 NITAAPADPILGLADLFRADERPGKINLIGIVYKDETKTPVLTSVKKAEOYLLLENETTK 63
Db 35 HVEMGPPDILGVTEAFKEDTNSKKNLGVGAYRDNGKPYVLPVSRKAEQAIAAKNLDK 94
Qy 64 NYLGIDGIDPEGRCTOELLFGKGSALINDKARTATQTPGGTGALRVAADFLAKNTSVKR- 122
Db 95 EYLPIGGLAEPCKAELALGENSEVLKSGREVTVTQTTSGTGALRIGASFQRPFFKSRD 154
Qy 123 VVVSNSPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSLEAQAQGDVVLPHGCC 182
Db 155 VFLPRTWGNHTPIFRDAGMQGGRYDYDPKCGDFGTGAVEDIISKIPEQSVLLHLHAC 214
Qy 183 NPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYOGFARG-LEBDAEGLRPAFAAMHKEIVA 241
Db 215 NPTGVDPRPEQNKIATVVVKKNLPAFPDMAYQGFASGDDGDADAVRHFIQGINVCLC 274
Qy 242 SSYSKNFGLYNERVGACTLVADSETVDRAFQSKAAIRANYSNPPAHGASVATILND 301
Db 275 QSYAKNMGLYGERVGAFTVCKDAEAKRVESQLKILIRPMYSNPPPLNGARITAAAILNTP 334
Qy 302 ALRATWEQBLTDMRORIORMROLFVNTLQEKANRDFSFIKQNGMFSFSGLTKEQVRL 361
Db 335 DURKQWLQEVKGMADRIIGMRTQLVSNLKSGSTWNHQTIDQIGMFCETGLKPEQVERL 394
Qy 362 RBEFGYAVASGRVNVAGMTPDNNAPLCEAI 392
Db 395 IKFEFSIYMTKDRISVAGVTSNVGYLAHI 425

RESULT 13

US-10-878-556A-149
; Sequence 149, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 149
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/aatc_human
; DATABASE ENTRY DATE: 1990-08-01
US-10-878-556A-149

Query Match 37.0%; Score 756; DB 6; Length 412;

Best Local Similarity 43.0%; Pred. No. 5.4e-56;
Matches 166; Conservative 54; Mismatches 156; Indels 10; Gaps 5;

Qy 10 ADPIL--GLADLFRADERPGKINLIGIVYKDETKTPVLTSVKKAEOYLL-ENETTKNYL 66
Db 12 AQPVLVFKLTADFPDPRKRVGLGVGAYRTDDCHPWLPVVKVQVEQKIANDNSLNHEYL 71
Qy 67 GIDGIDPEGRCTOELLFGKGSALINDKARTATQTPGGTGALRVAADFLAK-----NTSVK 121
Db 72 PILGLAEFRSCASRLALGDDSPALKEXRGVQSLGGTGALRICADFLARWYNGTNNKNT 131
Qy 122 RVVVSNSPNHKSFNVSAGL-EVREYAYDAENHTLDFDALINSLEAQAQGDVVLPHGC 180
Db 132 PVYVSSPTWENHNAVFAAGFKDIRSYRYMDAERKRGDLQGFLLNDLENAPESFVWLHAC 191
Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGWLPFLDFAYOGFARG-LEBDAEGLRPAFAAMHKEI 239
Db 192 AHNPTGIDPTPEQWKQIASVNMKHLFPFFDSATQGFASGNLERDANAIKRVFVSEGEFF 251
Qy 240 VASSYSKNFGLYNERVGACTLVADSETVDRAFQSKAAIRANYSNPPAHGASVATILS 299
Db 252 CAQSFSSKNFGLYNERVGNLTIVGKEPESILQVLSQMEKIVRIITWSNPPAQCARIVASTLS 311
Qy 300 NDALRAIWEQBLTDMRORIORMROLFVNTLQEKANRDFSFIKQNGMFSFSGLTKEQVL 359
Db 312 NPFLFEETWGNVKTMDRILTMRSRLARLEALKTPGTWNHITDQIGMFSFTGLNPKQVE 371
Qy 360 RLREEFQYAVASGRVNVAGMTPDNNM 385
Db 372 YLVNKEHIYLLPSGRINVSGLTTKNL 397

RESULT 14

US-11-245-400-17
; Sequence 17, Application US/11245400
; Publication No. US20060040357A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhkar
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Encoding Melanoma
; TITLE OF INVENTION: Associated Antigen Molecules, Aminotransferase
; TITLE OF INVENTION: Molecules, ATPase Molecules, Acyltransferase Molecules,
; TITLE OF INVENTION: Pyridoxal-Phosphate Dependant Enzyme Molecules and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/247400
; CURRENT APPLICATION NUMBER: US/11/245,400
; CURRENT FILING DATE: 2005-10-06
; PRIOR APPLICATION NUMBER: US/10/164,966
; PRIOR FILING DATE: 2002-06-07

```

; PRIOR APPLICATION NUMBER: 10/034,864
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/258,517
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/996,194
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/250,348
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,073
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/253,878
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,338
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 09/908,928
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/220,465
; PRIOR FILING DATE: 2000-07-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam Aminotransferase family model
US-11-245-400-17

Query Match      34.2%; Score 699; DB 7; Length 447;
Best Local Similarity 41.4%; Pred. No. 3.8e-51;
Matches 182; Conservative 60; Mismatches 152; Indels 46; Gaps 13;

Qy 1 MFENITAAPADPILGLADLFRADERPGKIN---LGIGVYKDETGTPTVLTSVKKAQOYL 56
Db 4 MAANVSHGPGPILGVGAFKEDPRGKDNPNVIGVAYEPQLGKDLVLPVAVKKAERKL 63
Qy 57 -----LENETTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGLARV 109
Db 64 ALDREGNIEFREIKYELPIHGLPEFREAIKLLFGARSPKLFKERVVVQTLGGTGLRL 123
Qy 110 AADFLAK---NTSVKRVVWVSNPNHKSVPNSAGLEV-REYAYY-DAENHTLDFDALI 163
Db 124 AADFLANPGDGRGSRGVLVPTWPNYKRDIFWAAAGVEVIVPYHYKDDENNFGLDFALE 183
Qy 164 NSLNEAQAGDV---VLFHGCCNPTGIDPTLEQWTLAQLSVEKGLPLDFPAYQGFARG 220
Db 184 AAIEKAPEKNITKTVLLHNNPHNPTGDTPTREQLKKIADVVKENILLSDAEYQGFVFG 243
Qy 221 -LEEDAEGLRAPAAHMK-----ELIVASSYSKNFGLYNERVGACTLV----- 261
Db 244 SLDEDAASVAFBAEVEKMECNGDELLLVQSFKNFGLYGRVGAIVVNPRTGDAVIS 303
Qy 262 -AADSETVDRAFSPQMAAIRANYSN---PPAHGASVATII-LSNDALRAIWEQELTDMRQR 317
Db 304 AAAXMSSAGRVSSQLQALARAMYSNPDPDPDHGAIEIVARILLERRDLFTSWLEEVKGMACR 363
Qy 318 IQMRM-QLFVNTLQKGNRDFSFIIKQNGHPSGLTKEQVLRIR--EEFGVAVASGR 374
Db 364 IPNGRLYLWMDLRKLLKEEDDWSHIIIEQEGMFSFTWLLNEEQVNVSPGSEPHIYEPGWR 423
Qy 375 VNVAGMTPDNNAPLCEAIVA 394
Db 424 ISLAGLSEANVEEAERIRA 443

```

RESULT 15

```

US-10-979-821-8
; Sequence 8, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.

```

```

; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: ZHAO, LISHAN
; APPLICANT: WEINER, DAVID P.
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 8
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Leishmania major
US-10-979-821-8

```

```

Query Match      30.9%; Score 632.5; DB 6; Length 412;
Best Local Similarity 38.0%; Pred. No. 1.3e-45;
Matches 150; Conservative 62; Mismatches 174; Indels 9; Gaps 6;

Qy 2 FENITAAPADPILGLADLFRADERPGKINLGIVYKDETGTPTVLTSVKKAQOYLLENET 61
Db 13 WQKIQAPQDVIFDLAKRAAAAGP-KANLVIGAYRDEQGRPYPLRVVRKABQLLLDMNL 71
Qy 62 TKNVYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGLARVAADFLAK-NTS 119
Db 72 DYEYLPISGYQPFIDEAVKIIYGNVLEEN---LVAVOTLSGTGAVSLGAKLLTRVFD 128
Qy 120 VKRVVWVSNPNHKSVPNSAGLE-VREYAYYDAENHTLDFDALINSLENAQAGDVVLF 178
Db 129 TPIYLSPTWPNHVGVAAGWKNICTYAYYDPKTVSLNFEKMKDILAAPDGSVFIH 188
Qy 179 GCCHNPTGIDPTLEQWTLAQLSVEKGLPLDFPAYQGFARG-LEEDAEGLRAPAAHMK 237
Db 189 QCAHNPTGVDPSPQSNBESIASMLAKHQVFFDSAYQGYASGSLDTDAYAARLFARRG 248
Qy 238 LIVASSYSKNFGLYNERVGACTLVAAADSETVDRAFSPQMAAIRANYSNPPAHGASV 297
Db 249 VLLAQSFKNMGLISERAGTSLLLKDKTKRADVKSVWDSLIRBEYTCPPAHGARLAH 308
Qy 298 LSNDALRAIWEQELTDMRQRIOQRMQLFVNTLQKGNRDFSFIIKQNGHPSGLTKREQ 357
Db 309 LSNNELRKWEAEELSAMAERIRTMRTVYDELLRLQTPGSMHVINOIGMFSFLGSKAQ 368
Qy 358 VLRUREFGVAVASGRVNVAGMTPDNNAPLCEAI 392
Db 369 -CEYCONHNIPITVSGRANMAGLTHTALMLAQTI 402

```

Search completed: March 14, 2006, 18:15:20
Job time : 24 secs

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 14:05:42 ; Search time 5402 Seconds
(without alignments)
10315.330 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191

Sequence: 1 atgtttgagaacattaccgc.....cgattgtgcagtgctgtaa 1191

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hc: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gss1: *
10: gb_gss2: *
11: gb_gss3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	855	71.8	864	10	CL660432
2	783.4	65.8	785	10	CL695845
3	563	47.3	671	10	CL695172
4	556.8	46.8	645	6	CA093007
5	482.4	40.5	557	6	CF306417
6	445.8	37.4	468	1	AW221110
7	236.8	19.9	1690	4	AT103658
8	228.8	19.2	1749	8	CK109239
9	224.8	18.9	1832	4	AY105250
10	220.4	18.5	2452	4	CK857622
11	215	18.1	516	6	CF338891
12	211.6	17.8	2447	4	CF3860094
13	205.4	17.2	1599	4	CNS0A2EC
14	203	17.0	977	7	CK259742
15	200.6	16.8	1486	4	CNS092WQ
16	199.6	16.8	902	7	CK259741
17	190.4	16.0	1643	4	CNS0A80C
18	190	16.0	858	7	CO101379
19	189.6	15.9	919	7	CO732049
20	188.2	15.8	968	7	CK262759
21	188	15.8	704	3	BJ624279
22	187	15.7	914	7	CO732002

23	185.6	15.6	852	7	CO459606	CO459606	MZCCL2002
24	185.6	15.6	876	8	DR059430	DR059430	RTNITI 17
25	184.2	15.5	759	6	CD919780	CD919780	G608.114J
26	183	15.4	854	7	CO464147	CO464147	MZCCL1504
27	182.6	15.3	783	8	DN147062	DN147062	4853 All
28	182.6	15.3	1564	4	CNS0EPVU	CNS0EPVU	CR636228 Tetraodon
29	179.8	15.1	847	6	CD574536	CD574536	UCRPT01_0
30	179.4	15.1	2017	4	CR860603	CR860603	Pongo pyg
31	179	15.0	1131	7	CV734832	CV734832	CR12 11 H
32	179	15.0	1440	4	CNS0BOL3	CR634545	Tetraodon
33	177.8	14.9	1496	4	CR616132	CR616132	full-leng
34	177.8	14.9	1509	4	CR628608	CR628608	full-leng
35	177.8	14.9	1512	4	CR605482	CR605482	full-leng
36	177.8	14.9	1525	4	CR615083	CR615083	full-leng
37	177.8	14.9	1540	4	CR598904	CR598904	full-leng
38	177.8	14.9	1560	4	CR622320	CR622320	full-leng
39	177.8	14.9	1563	4	CR601723	CR601723	full-leng
40	177.8	14.9	1719	4	CR606661	CR606661	full-leng
41	177.8	14.9	1959	4	CR625122	CR625122	full-leng
42	177.8	14.9	1979	4	CR619243	CR619243	full-leng
43	177.8	14.9	1979	4	CR620721	CR620721	full-leng
44	177.6	14.9	1030	9	BZ567307	BZ567307	paca2-164
45	176.6	14.8	1517	4	CNS0GMY6	CR725768	Tetraodon

ALIGNMENTS

RESULT 1
CL660432
LOCUS
DEFINITION
CL660432 864 bp DNA linear GSS 09-JUL-2004
PRI0137a.D10 - PRI0137a.B21 (864) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
CL660432
CL660432.1 GI:50145720
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
1 (bases 1 to 864)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an AcedB database for the nematode satellite organism Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
Location/Qualifiers
1..864
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pBifos-5 Fosmid vector"

ORIGIN

Query Match 71.8%; Score 855; DB 10; Length 864;
Best Local Similarity 100.0%; Pred. No. 9.4e-257;
Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 55 CTGTTTCGCGGTACGATCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGAT 114
Db 10 CTGTTTCGCGGTACGATCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGAT 69
Qy 115 GAGACGGCAAAACCCCGGTACTGACACGCGTGAAAAAGGCTGAACAGTATCTCTCGAA 174
Db 70 GAGACGGCAAAACCCCGGTACTGACACGCGTGAAAAAGGCTGAACAGTATCTCTCGAA 129
Qy 175 AATGAAACACCAAAATTAACCTCGGCATTTGACGCGATCCCTGAATTTGGTGTCTGACT 234
Db 130 AATGAAACACCAAAATTAACCTCGGCATTTGACGCGATCCCTGAATTTGGTGTCTGACT 189
Qy 235 CAGGAACCTGCTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCGCACGCA 294
Db 190 CAGGAACCTGCTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTCGCACGCA 249
Qy 295 CAGACTCCGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACC 354
Db 250 CAGACTCCGGGGGACCTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACC 309
Qy 355 AGCGTTAAGCTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGGCTCTTTAAC 414
Db 310 AGCGTTAAGCTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGGCTCTTTAAC 369
Qy 415 TCTGAGGCTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTAC 474
Db 370 TCTGAGGCTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATCACACTCTTAC 429
Qy 475 TTCATGACCTGATTAAACAGCTCTGAATGAAGCTCAGGCTGGCGAGCTAGTCTCTTCCAT 534
Db 430 TTCATGACCTGATTAAACAGCTCTGAATGAAGCTCAGGCTGGCGAGCTAGTCTCTTCCAT 489
Qy 535 GGCTGCTGCCATAACCCAAACCGGTATGACCTTACGCTTGGAAACCAATGGCAAAACCTGCA 594
Db 490 GGCTGCTGCCATAACCCAAACCGGTATGACCTTACGCTTGGAAACCAATGGCAAAACCTGCA 549
Qy 595 CAACCTCCGTTGAGAAAGGCTGGTACCGCTGTTGACTTTCGCTTACCAGGTTTGGC 654
Db 550 CAACCTCCGTTGAGAAAGGCTGGTACCGCTGTTGACTTTCGCTTACCAGGTTTGGC 609
Qy 655 CGTGTCTGGAAGAGATGCTGAAGGACTGGCGCTTTTCGCGGCTATGCATAAAGAGCTG 714
Db 610 CGTGTCTGGAAGAGATGCTGAAGGACTGGCGCTTTTCGCGGCTATGCATAAAGAGCTG 669
Qy 715 ATTGTTCAGTTCCTACTATAAACTTTGGCTGTATACACGAGCGTGTGGCGCTTGT 774
Db 670 ATTGTTCAGTTCCTACTATAAACTTTGGCTGTATACACGAGCGTGTGGCGCTTGT 729
Qy 775 ACTCTGTTGCTGCCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCG 834
Db 730 ACTCTGTTGCTGCCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGCG 789
Qy 835 ATTCCGCTTAACCTTAACCCAGCAGCAGCGCTTCTGTTGTTGCCACCATCCTG 894
Db 790 ATTCCGCTTAACCTTAACCCAGCAGCAGCGCTTCTGTTGTTGCCACCATCCTG 849
Qy 895 AGCAACGATGCGTTA 909
Db 850 AGCAACGATGCGTTA 864

```

```

RESULT 2
LOCUS CL695845
DEFINITION PRI017b H12.2 - PRI017b.BR (785) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL695845
VERSION CL695845.1 GI:50217753
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

```

```

Neodiplogasteridae; Pristionchus.
1 (bases 1 to 785)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES             Location/Qualifiers
     source            1..785
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN
Query Match      65.8%; Score 783.4; DB 10; Length 785;
Best Local Similarity 99.9%; Pred. No. 3e-234;
Matches 784; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 380 ACCCAAGCTGCGCAACCATTAAGAGCGCTTTTAACTCTGCAAGTCTGGAAGTTGCTGAAT 439
Db 1 ACCCAAGCTGCGCAACCATTAAGAGCGCTTTTAACTCTGCAAGTCTGGAAGTTGCTGAAT 60
Qy 440 ACCTTATTATGATGCGGAAATCACACTCTTCACTTCGATGCACTGATTAAACAGCTGA 499
Db 61 ACCTTATTATGATGCGGAAATCACACTCTTCACTTCGATGCACTGATTAAACAGCTGA 120
Qy 500 ATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGTCGCAATACCCACCGGTA 559
Db 121 ATGAAGCTCAGGCTGGCGAGCTAGTGTCTTCCATGGCTGTCGCAATACCCACCGGTA 180
Qy 560 TCGACCTCAGCTGGAAACAATGGCAACAACCTGCGTTCGTTGAGAAGGCTGGT 619
Db 181 TCGACCTCAGCTGGAAACAATGGCAACAACCTGCGTTCGTTGAGAAGGCTGGT 240
Qy 620 TACCGCTGTTGACTTCGCTTACAGGGTTTTCGCGGTTTTCGCGGTTGCTGGAAGAGATGCTGAAG 679
Db 241 TACCGCTGTTGACTTCGCTTACAGGGTTTTCGCGGTTTTCGCGGTTGCTGGAAGAGATGCTGAAG 300
Qy 680 GACTGCGGCTTTTCGCGGCTATGCATAAAGAGCTGATTGTTGCCAGTTCTTACTCTAAAA 739
Db 301 GACTGCGGCTTTTCGCGGCTATGCATAAAGAGCTGATTGTTGCCAGTTCTTACTCTAAAA 360
Qy 740 ACTTTGGCTGTACAACGAGCGTGTGGCGCTTGTACTCTGTGTTGCTGCGAGAGTGA 799
Db 361 ACTTTGGCTGTACAACGAGCGTGTGGCGCTTGTACTCTGTGTTGCTGCGAGAGTGA 420
Qy 800 CCGTTGATCGGCATTCAGCCAAATGAAGGGGATTCGCGCTAACTACTCTAACCCAC 859
Db 421 CCGTTGATCGGCATTCAGCCAAATGAAGGGGATTCGCGCTAACTACTCTAACCCAC 480
Qy 860 CAGCACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAACGATGCGTTCAGTCCGATTT 919
Db 481 CAGCACACGCGCTTCTGTTGTTGCCACCATCTCGAGCAACGATGCGTTCAGTCCGATTT 540
Qy 920 GGGAAACAAGAGCTGACTGATATGCGCAGCGGTATTCAGCGTATGCGTTCAGTTGTTGCTCA 979
Db 541 GGGAAACAAGAGCTGACTGATATGCGCAGCGGTATTCAGCGTATGCGTTCAGTTGTTGCTCA 600

```


temperature stress). cDNA was prepared from poly(A⁺ mRNA using ZAP - cDNA Synthesis kit (Stratagene). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public/>

ORIGIN

Query Match	46.8%;	Score 556.8;	DB 6;	Length 645;
Best Local Similarity	99.6%;	Pred. No. 4.9e-163;		
Matches 558;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	90	CCTCGGATTGGTGTCTATAAAGATGAGCGGCAAAACCCCGGTACTGACCCAGCGTGA	149	
Db	21	CCTCGGATTGGTGTCTATAAAGATGAGCGGCAAAACCCCGGTACTGACCCAGCGTGA	80	
Qy	150	AAAGGCTGAACAGTATCTGCTCGAAATGAACACCAAAATTAATCTCGGCATTTGACGG	209	
Db	81	AAAGGCTGAACAGTATCTGCTCGAAATGAACACCAAAATTAATCTCGGCATTTGACGG	140	
Qy	210	CATCCCTGAATTTGGTGCCTGCACCTCAGGAACCTGCTTTGGTAAAGGTAGCGCCCTGAT	269	
Db	141	CATCCCTGAATTTGGTGCCTGCACCTCAGGAACCTGCTTTGGTAAAGGTAGCGCCCTGAT	200	
Qy	270	CAATGACAAACGCTGCTCGCAGCGCACAGACTCCGGGGGGCACCTGGCGCACTACGCGTGGC	329	
Db	201	CAATGACAAACGCTGCTCGCAGCGCACAGACTCCGGGGGGCACCTGGCGCACTACGCGTGGC	260	
Qy	330	TGCCGATTTCTTGCAAAAAAATACGAGTTAAGCGTTAGCGTGTGGGTGAGCAACCCAGCTG	389	
Db	261	TGCCGATTTCTTGCAAAAAAATACGAGTTAAGCGTGTGGGTGAGCAACCCAGCTG	320	
Qy	390	GCCGAACCAATGAAGCGCTTTAACTCTGCAGGCTCGGAAGTTCGTGAATACGCTTATTA	449	
Db	321	GCCGAACCAATGAAGCGCTTTAACTCTGCAGGCTCGGAAGTTCGTGAATACGCTTATTA	380	
Qy	450	TGATGCCGAAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCCTGAATGAAGCTCA	509	
Db	381	TGATGCCGAAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCCTGAATGAAGCTCA	440	
Qy	510	GGCTGGCGACGTAGTGTCTTCCATGGCTGCTGCCATAACCCAAACCGGTATCGACCCCTAC	569	
Db	441	TGCTGGGCGACGTAGTGTCTTCCATGGCTGCTGCCATAACCCAAACCGGTATCGACCCCTAC	500	
Qy	570	GCTGGAACAATGGCAAAACACTGGCAACAATCTCCGTTTGAGAAAGCGTGGTTACCGCTGTT	629	
Db	501	GCTGGAACAATGGCAAAACACTGGCAACAATCTCCGTTTGAGAAAGCGTGGTTACCGCTGTT	560	
Qy	630	TGACTTCGCTTACCAGGTT	649	
Db	561	TGACTTCGCTTACCAGGTT	580	

RESULT 5	CF306417	CF306417	LOCUS	CF306417	557 bp	mRNA	linear	EST 15-AUG-2003
DEFINITION	HDAL1--03-M15.g1 OsHDA1-overexpressing transgenic rice lambda phage cDNA library 1 (HDAL1) Oryza sativa (japonica cultivar-group) cDNA clone HDAL1--03-M15, mRNA sequence.							

ACCESSION

ACCESSION
VERSION

NOTES
SUBWAYS

KEYWORDS

SOURCE
ORGANIC

ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

REFERENCES

Overview

TITLE.

RESIST. 6

KEY: 0
AW221110

LOCUS
T222110

LOCUS
DEFINITION

DEFINITION

Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhna@bio.myongji.ac.kr.

FEATURES
source

```

i: 351
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="HDA1--03-M15"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 2 weeks"
/clone_lib="OshDAC1-overexpressing transgenic rice lambda
phage cDNA library I (HDA1)"
/notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; Callus was treated with ABA(20um) for 1hour. cDNA
was inserted into lambda Uni-XR vector at 5' end with
EcoRI and 3' end with XhoI site. mRNA was derived from
rice Histone Desacetylase overexpression line."

```

ORIGIN

Query Match	40.5%;	Score 482.4;	DB 6;	Length 557;
Best Local Similarity	99.8%;	Pred. No. 1.1e-139;		
Matches 483;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	505	GCTCAGGCTGCGCAGCTAGTGTCTGTCATCTGGCTGCTGCCATAACCCAAACCGGTATCGAC	564	
Db	74	GCCCAGGCTGCGCAGCTAGTGTCTGTCATGGCTGCTGCCATTAACCCNACCGGTATCGAC	133	
QY	565	CCTACGCTGGAAACAATGGCAACAACCTGCGTTCGTTGAGAAAGGCTGGTTACCG	624	
Db	134	CCTACGCTGGAAACAATGGCAACAACCTGCGTTCGTTGAGAAAGGCTGGTTACCG	193	
QY	625	CTGTTTGACTTCGCTTACACAGGGTTTTGCCGCTGGTCTGGAAGAAGATGCTGAAGCACTG	684	
Db	194	CTGTTTGACTTCGCTTACACAGGGTTTTGCCGCTGGTCTGGAAGAAGATGCTGAAGCACTG	253	
QY	685	CGCGCTTTTCGGCGCTATGTCATAAAGAGAGTGATTGTTGCCAGTTCCTACTCTAAAAA	744	
Db	254	CGCGCTTTTCGGCGCTATGTCATAAAGAGAGTGATTGTTGCCAGTTCCTACTCTAAAAA	313	
QY	745	GGCCTGTACACAGCAGCGTGTGGCGCTTGTA	804	
Db	314	GGCCTGTACACAGCAGCGTGTGGCGCTTGTA	373	
QY	805	GATCGCGATTTCAGGCAATGAAGACGGCGGATTCGCGCTAACTACTCTAACCCACCAGCA	864	
Db	374	GATCGCGATTTCAGGCAATGAAGACGGCGGATTCGCGCTAACTACTCTAACCCACCAGCA	433	
QY	865	CACGCGCTCTCTGTTGTTGCCACCATCTCAGCAACGATGCGTTACGTGCGATTTCGGAA	924	
Db	434	CACGCGCTCTCTGTTGTTGCCACCATCTCAGCAACGATGCGTTACGTGCGATTTCGGAA	493	
QY	925	CAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGGTATGCGTCAGTGTTCGTCATATCG	984	
Db	494	CAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGGTATGCGTCAGTGTTCGTCATATCG		
QY	985	CTGC 988		
Db	554	CTGC 557		

AW221110 468 bp mRNA linear EST197579 tomato fruit mature green, TAMU Lycopersicon esculentum cDNA clone cLEF12H11, mRNA sequence.
AW221110


```
VERSION AW221110.1 GI:6532794
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 468)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., and
Roning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D., and
Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source Location/Qualifiers
1..468
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF12H11"
/tissue_type="fruit pericarp"
/dev_stages="mature green (3-5 days pre-ripening)"
/lab_host="SOUR"
/clone_lib="tomato fruit mature green, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
ORIGIN
Query Match 37.4%; Score 445.8; DB 1; Length 468;
Best Local Similarity 99.6%; Pred. No. 3.5e-128;
Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 153 GCCTGAACAGTATCTGCTCGAAATGAAACCAACCAAAATTAACCTCGGCATTGACGGCAT 212
Db 1 GCCTGAACAGTATCTGCTCGAAATGAAACCAACCAAAATTAACCTCGGCATTGACGGCAT 60
Qy 213 CCCTGAATTTGGTGCCTGCACTCAGGAATCTGCTGTTTGGTAAAGTAGCGCCCTGATCAA 272
Db 61 CCCTGAATTTGGTGCCTGCACTCAGGAATCTGCTGTTTGGTAAAGTAGCGCCCTGATCAA 120
Qy 273 TGACAAAGCTGCTCGACGGCACAGACTCGGGGGGCACTCGGCAGTACGCGTGGCTGC 332
Db 121 TGACAAAGCTGCTCGACGGCACAGACTCGGGGGGCACTCGGCAGTACGCGTGGCTGC 180
Qy 333 CGATTTCTCGCAAAATACACAGCTTAAGCGTGTGGGTGAGCAACCAAGCTGCC 392
Db 181 CGATTTCTCGCAAAATACACAGCTTAAGCGTGTGGGTGAGCAACCAAGCTGCC 240
Qy 393 GAACCATAGAGCGCTTTAACTCTGCAGGCTCTGGAAGTTCGTGAATACGCTTATTATGA 452
Db 241 GAACCATAGAGCGCTTTAACTCTGCAGGCTCTGGAAGTTCGTGAATACGCTTATTATGA 300
Qy 453 TGGCGAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCTGAATGAAGCTCAGGC 512
Db 301 TGGCGAAAATCACACTCTTGACTTCGATGCACTGATTAACAGCCTGAATGAAGCTCAGGC 360
Qy 513 TGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
Db 361 TGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 573 GGAAACAATGGCAACACTGGCACAACTCT 601
```

```
Db 421 GGAAACAATGGCAACACTGGCACAACTAT 449
RESULT 7
AY103658 1690 bp mRNA linear HTC 18-FEB-2005
LOCUS Zea mays PCO118382 mRNA sequence.
DEFINITION Zea mays PCO118382 mRNA sequence.
ACCESSION AY103658
VERSION AV103658.1 GI:21206736
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1690)
AUTHORS Gardiner,J., Schroeder,S., Polacco,M.L., Sanchez-Villeda,H.,
Fang,Z., Morgante,M., Landewe,T., Fengler,K., Useche,F.,
Hanafey,M., Tingey,S., Chou,H., Wing,R., Soderlund,C. and Coe,E.H.
Jr.
TITLE Anchoring 9,371 maize expressed sequence tagged unigenes to the
bacterial artificial chromosome contig map by two-dimensional
overgo hybridization
JOURNAL Plant Physiol. 134 (4), 1317-1326 (2004)
PUBMED 15020742
REFERENCE 2 (bases 1 to 1690)
AUTHORS Haney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 3 (bases 1 to 1690)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org: ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source Location/Qualifiers
1..1690
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeGDB:636895"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
ORIGIN
Query Match 19.9%; Score 236.8; DB 4; Length 1690;
Best Local Similarity 51.5%; Pred. No. 2.9e-62;
Matches 597; Conservative 0; Mismatches 557; Indels 6; Gaps 2;
Qy 22 GTCCTCGCCAGCCCGATTCGGCGCTCGCGCATCTGTTTGTGCGGATGAACGTCGCCGC 81
Db 264 GCCCGGAGGACCCCATCTCGGAGTGACCGTTGCCCTACAAAGGATCCCGAGCCCGTG 323
Qy 82 AAAATTAACTCGGATTTGCTATTAAGATGAGACGGCAAAACCCCGGTACTGACC 141
Db 324 AAGTCAACTTCGGGGTTCGGCGCTACCGGACCGGAGGAGGAGCCCTAGTCTGAAC 383
Qy 142 AGCGTAAAAAGGCTGAACAGTATCTCTCGAAATGAACACAC- - -CAAAAATTACCTC 198
```

LOCUS	CX109239	1749 bp	mRNA	linear	EST 03-JUN-2005
DEFINITION	RECM0798 A normalized whole-life-cycle cDNA library of rice Oryza sativa (indica cultivar-group) cDNA clone B1130A13, E1078H06, B1100J15, B1107E19, B1131F09, B1075E21, B107 5', mRNA sequence.				
ACCESSION	CX109239				
VERSION	CX109239.1				
KEYWORDS	EST.				
SOURCE	Oryza sativa (indica cultivar-group)				
ORGANISM	Oryza sativa (indica cultivar-group)				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.				
AUTHORS	1. (bases 1 to 1749) Zhang,J., Feng,Q., Jin,C., Qiu,D., Zhang,L., Xie,K., Yuan,D., Han,B., Zhang,Q. and Wang,S.				
TITLE	Features of the expressed sequences revealed by a large-scale analysis of ESTs from a normalized cDNA library of the elite indica rice cultivar Minghui 63				
JOURNAL	Plant J. 42 (5), 772-780 (2005)				
PUBMED	15918889				
COMMENT	Contact: Wang S National Key Laboratory of Crop Genetic Improvement Hauzhong Agricultural University Wuhan 430070, China Tel: 86-27-87282044 Fax: 86-27-87287092 Email: shiyingwang@hotmail.com Seq primer: T7.				
FEATURES	Location/Qualifiers				
source	1..1749 /organism="Oryza sativa (indica cultivar-group)" /mol_type="mRNA" /strain="indica" /cultivar="Minghui 63" /db_xref="taxon:39946" /clones="B1130A13, E1078H06, B1100J15, B1107E19, B1131F09, B1075E21, B107" /tissue_type="whole plant" /dev_stage="whole-life-cycle" /lab_host="E. coli DH10B" /clone_lib="A normalized whole-life-cycle cDNA library of rice" /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; The library is constructed based on the strategy of saturation hybridization with genomic DNA using rice cultivar Minghui 63. This library consists of cDNA from 15 directionally cloned cDNA libraries constructed with different tissues from 9 developmental stages."				
ORIGIN	Query Match 19.2%; Score 228.8; DB 8; Length 1749; Best Local Similarity 51.0%; Pred. No. 9.6e-60; Matches 592; Conservative 0; Mismatches 562; Indels 6; Gaps 2;				
Qy	22	GCTCCTGCGACCGGATTCCTGGGCTCGCGCATCTGTTTCTGTCGGCATGAACGTCCTCCGCG	81		
Db	302	CGCGCGGAGGACCCCATCTCGGGGTACCGTCCGCTACACAAGGATCCCGGCGCGGTG	361		
Qy	82	AAATTAACCTCGGGATTTGTTGTTCTATAAAGATGAGACGGGCAAAACCCCGGTACTGACC	141		
Db	362	AAGGTCAATCTCGCGTCGCGGCTACCGGACCGGAGGAGGAGCCCTGCTGCTGAAT	421		
Qy	142	AGCGTGAAGGCTGAACAGTATCTGCTCGAATAATGAACACCACCC---AAAAATTACCTC	198		
Db	422	GTGTCAGCGCGCTGAGCAGATGCTGATCAACAACCCGTCACGTTTAAAGGATATTG	481		
Qy	199	GGCATTTGACGCGATCCCTCGAATTTGGTCGCTCAGGAACCTGCTGTTTGTAAAGGT	258		
Db	482	CCGATCACTGGACTGGCCGATTTCAATAAGCTGAGTCTAAGCTTATTTTGGTGTGCTGAC	541		
Qy	259	AGCGCCCTGATCAATGACAAACGTCCTCGCACGGCACAGACTCCGGGGGCACTGGCGCA	318		
Db	542	AGTCTGCCATTCAAGAGATAGGGTGGCTACAGTTCACTGCTTGTTCAGGAACCTGGTTCT	601		

Qy 319 CTACGGTGGCTCCGATTTCTTGGCAAAAAATACCAGCGTTAAGCGTGTGGGTGAGC 378
Db |||||
602 TTAAGGGTGGAGGTGAATTTCTTCAAGGCATTATCATGAAGCGCTATATACATCCCA 661
Qy 379 AACCAAGCTGCCCAACCATGAAGAGCTCTTTAACTCTGAGGCTGGAAGTTGTTGAA 438
Db |||||
662 CAGCCAACTGGGGAAATCACCCAAAGTGTTCACCTTTAGCTGGCTGACTGTTAGGAGT 721
Qy 439 TACGGCTTATATGATGCGGAAATCACACTCTTGACTTCGATGCACTGATTAACAGCGTG 498
Db |||||
722 TACCGCTACTATGATCTTCAACCCGTGAGCTGGATTTCCAAAGGCTGTAGAAGATCTC 781
Qy 499 AATGAAGCTCAGGCTGGCGAAGTAGTGTGTTTCCATGCTGTGCGATTAACCCACCGGT 558
Db |||||
782 GGTTCAGCTCTTCCAGGTGCAATGTTACTGCTTCATGCTTGTGCGCCACAAACCTACTGA 841
Qy 559 ATCAGACCTAGCTGGAACATGGCAACACATGGCACAACTCCGTTGAGAAAGCGTGG 618
Db |||||
842 GTAGACCAACTTTGGACGAGTGGGAACAGATCAGGCAGTTGATGAGATCAAAAGCATTG 901
Qy 619 TTACCGCTGTTGACTTCGCTTACCAGGTTTTGCGCGTG---GTCGGAAGAGATGCT 675
Db |||||
902 CTGCCATCTTTGATAGCGCTTATCAGGGAATTTGCAAGTGAAGTCTTGACCAAGATGCT 961
Qy 676 GAAGGACTCGCGCTTTTCGCGCTATGCATGAAGAGCTGATTTGCGCAGTTCTCTACTCT 735
Db |||||
962 CAATCAGTGGCGATGTTGTTGCTGATGTTGTTGCTGATGCTCAGAGCTACGCT 1021
Qy 736 AAAAATTGCGCTGTACAACAGAGGTTGCGGCTGTGCTGCTGCTGCGGACAGT 795
Db |||||
1022 AAGAACATGGGATTTGATGGAGAGCGTGTGCGTCTTTAAGCATATGTTTGTGGAAGTCT 1081
Qy 796 GAAACCGTGTATCGCGCATTCAGCCAAATGAAGCGCGATTCGCGCTAACTACTCTAAC 855
Db |||||
1082 GATGAGCTGTGAGGTTGAAAGTCAACTTAAGCTTGTAAATAGGCTGATGATTCAAAAC 1141
Qy 856 CCACGAGCACGCGGCTTCTGTTGTCACCATCTTGAGCAAGCATGCGTTACGTGCG 915
Db |||||
1142 CCTCTATTATGTTGTCATCTATCTGCTGCTACCATCTTAAGGACAGTGAATGTTCAAT 1201
Qy 916 ATTTGGGAACAGAGCTGATGATGCGCCAGCGTATTCAGCGTATGCGTCAAGTTGTC 975
Db |||||
1202 GAATGGAGCTGTGGAGCTGAAGGCGATGCTGATAGGATTAATAGCATGAGGCAACAGCTT 1261
Qy 976 GTCATACGCTGCAGGAAGGCGCAACACCGGACTTCAGCTTTATCATCAACAGAC 1035
Db |||||
1262 TTTGATGCTCTGAAACAGAGAAACTCCTGAGAGCTGGAGTCAATCATTAAGCAGATC 1321
Qy 1036 GGCATGTTCTCTTTCAGTGGCTGTACAAAAGAAAGTGTGCGTCTGCGGAAGAGTTT 1095
Db |||||
1322 GGAATGTTCACTTTTCACTGGGCTCAACAGTATCAAGTGGCTTTCATGAGGCGAGATAC 1381
Qy 1096 GCGGTATATGCGGTTGCTTCTGTCGCGTAAATGTGGCCGGAATGACACAGATTAACATG 1155
Db |||||
1382 CACATTTACATGACATCTGATGGGAGGATCAGTATGGCTGTTGAGCGGGAGGACCATT 1441
Qy 1156 GCTCGGCTGTGGAGCGCAT 1175
Db |||||
1442 CCGCATCTTGAGATGCAAT 1461

RESULT 9
AY105250
LOCUS Zea mays PCO107699 mRNA sequence. 1832 bp linear HTC 18-FEB-2005
DEFINITION Zea mays PCO107699 mRNA sequence.
ACCESSION AY105250
VERSION AY105250.1 GI:21208328
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE
AUTHORS

TITLE

JOURNAL
PUBMED

REFERENCE
AUTHORS

TITLE

JOURNAL
AUTHORS

JOURNAL

COMMENT

FEATURES
source

clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1832)
Gardiner, J., Schroeder, S., Polacco, M.L., Sanchez-Villeda, H.,
Fang, Z., Morgante, M., Landewe, T., Fengler, K., Useche, F.,
Hanafey, M., Tingey, S., Chou, H., Wing, R., Soderlund, C. and Coe, E.H.
Jr.
Anchoring 9,371 maize expressed sequence tagged unigenes to the
bacterial artificial chromosome contig map by two-dimensional
overgo hybridization
Plant Physiol. 134 (4), 1317-1326 (2004)
15020742
2 (bases 1 to 1832)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
3 (bases 1 to 1832)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MST, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1..1832
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:636247"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 18 9%; Score 224.8; DB 4; Length 1832;
Best Local Similarity 51.1%; Pred. No. 1.8e-58;
Matches 605; Conservative 0; Mismatches 572; Indels 7; Gaps 3;

Qy 4 TTTGAGAACATTACCGCGCTCTGCGGACCCGATTTCTGGGCTGCGCGATCTGTTTGGT 63
Db |||||
386 TTTGAGAGGTGCCAATGGCTCTCCAGACCCCAATCTTGGGTTCTGAGGCTTTAA 445
Qy 64 GCCGATGAACGTCCTCCGCAAAATTAACCTCGGGAATGGTGTCTATAAAGATGAGACGGC 123
Db |||||
446 GCAGATAAAGCGAGCTGAAGCTCAATCTTGGTGTGGTGCCTATAGGACAGAGAGCTG 505
Qy 124 AAAACCCCGGTACTGACCGGTCGAAAAGCTGAAAGATATCTGCTCGAANAATGAACCC 183
Db |||||
506 CAGCCCTACGTGCTCAATGTAGTCAAGAAGCTGAAAATCTTATGTGTGAGAAAGGAGAA 565
Qy 184 ACCAAAAATTACCTCGGCATTGACGGCATCCCTGAAATTTGGTCGTCGACCTCAGGAACCTG 243
Db |||||
566 AACAAAGAGTATCTTCCCATTTGAAGTTTAGCCGGTTTAAACAAGCAACAGCAGAGTA 625
Qy 244 CTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAAGTGTCTCGACGGCAGACACTCCG 303
Db |||||
626 TTGCTTGGAGCTGATAACCCCTGTTTAAATCAAGGACTGGTTGCTACACTTCAGTCTCTC 685
Qy 304 GGGGGGACTGGCGCACTACGGTGGCTGCCGATTTCTTGGCAAAAATACCAGCGTTAAG 363
Db |||||
686 TCGGGGCACTGGATCACTCGCTCGCTGCAGCATTTATCAAAAGATACTTTTCTCTGAAGCT 745

Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?ClonedID=DKFZp469B0125>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

source
1. .2447
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469B0125"
/tissue_type="kidney"
/clone_lib="469 (synonym: pkidi). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiIb"
/dev_stages="adult"
/note="aspartate transaminase precursor, mitochondrial
(Homo sapiens)"
1. .2447
/gene="DKFZp469B0125"
join(103..1116,1116..1394)
/gene="DKFZp469B0125"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH92240.1"
/db_xref="GI:55731048"
/translation="MALLHSGRLPGIAAPHPCGLAAASARASSWTHVEMGPPDI
LGVTAEKFRDINSKMLGVGYRDDNGKPYVLPSVRKAEQIAKRLDKEYLPIGL
AEFCASAEALNSSEVLKSGRFVTQISGTALRIKIGASFLQRFKSRDVPFLPKP
SWGHTIFRDAGVOLQGYRYDPKTFDFTGAVEDI SKIPEOSVLLHACAHNPFG
VDPRPEWKETVKKRNLPFAFDMAYQGFASGDGDKDAWVRHFTTEQGINVCLCS
YAKNMGILGERVGAFTWVCKDAEKVESOLKILIRPMYSNPPLNGARIAAAILNP
DLRTWLEQVKGMDRLIIGMRTOLVSNLKKEGSTHWHITDQIGMFCFTGLRPEQVE
RLIKEFSIYMTKGRISVAGVTSSNVGLAHAIHQVTK"

ORIGIN

Query Match 17.8%; Score 211.6; DB 4; Length 2447;
Best Local Similarity 50.7%; Pred. No. 2.8e-54;
Matches 587; Conservative 0; Mismatches 564; Indels 7; Gaps 3;
25 C T G C C G A C C C G A T T C T G G G C C T G G C C G A T C T G T T C G T C C G A T G A A C C C C G G C A A A 84
220 C C T C C A G A T C C A T T C T G G A G T C A C T G A A G C C T T T A A G A G G A C A C C A A T A G A A A A A G 279
85 A T T A A C C T C G G A T T G G T G T C T A T A A G A T G A G A C G G C A A A C C C C G G T A C T G A C C A G C 144
280 A T G A A T C T G G A G T T G G T G C T A C C G G A T G A T A C G G A A A G C C T T A C T G C T G C C T A G C 339
145 G T G A A A A G G C T G A A C A G T A T C T G C T C G A A A T G A A A C C A C C A A A A A T T A C T C G G C A T T 204
340 G T C C C A G G C A G A G C C C A G A T T C C C G C A A A A A T T T G G A C A G G A T A C T C G C C C A T T 399
205 G A C G G C A T C C C T G A A T T T G G T G C G T G C A C T C A G G A A C T G C T G T T T G T T G A A G G T A G C G C C 264
400 G G G G A C T G G T G A A T T T T G C A A G G C A T C T C A G A A C T A G C C C T G G G T G A G A A C A G C G A A 459
265 C T G A T C A T G A C A A C G T C T C G C A G G C A C A G A C T C G G G G G C A C T G G G C A C T A C G C 324
460 G T C T T G A A G A G T G G C G G T T G T C A C T G T G C A G A C C A T T T C T G G A A C T G S G A C C T T A A G G 519
325 G T G G T C G C G A T T T C C T G C G A A A A A T A C C A G C G T T A A G C G - - - T G T G G G T G A G C A A C 381
520 A T C G A G C A G T T T T C T G C A A G A T T T T T A A G T T C A C C C A G A T G T T T T C G C C C A A 579
382 C C A A G C T G C C G C A A C C A T A A G A G C G T C T T T A A C T C T G C A G T C T G G A A G T T C G T G A T A C 441
580 C C A T C C T G G G A A C C A C A C C C A C C A T T C T C A G G A T G T G G A T C G C A G T C A C A A G G T T A T 639
442 G C T T A T T A T G A T G C G G A A A A T C A C A C T C T T G A C T T C G A T C G A T C A T T A A C A G C C T G A T 501
640 C G G T A T T A T G A C C C A A G A C T T G C G G T T T T G A C T T C A C A G T G T G T G G A G A T A T T T C A 699
502 G A A G C T A G G C T G G C G A C T A G T G T G T T C A T G C T G C T A A C C A A C C A A C C G G T A T C 561

700 A A A A T A C C A G A G A G A G T G T T C T T C T T G C A T G C C T G C G C C C A C A A T C C C A C G G A G T G 759
562 G A C C T A C G C T G A A C A A T G G C A A C A C T C G C A C A A C T C C G T T G A G A A A G G C T G T T A 621
760 G A C C C G C T C C G A A C A G T G A A G A A A T A G A A C A C A G T G G T G A A A A A A A A G A A T C T C T T T 819
622 C C G C T G T T G A C T T C G C T T A C C A G G G T T T T G C C C G T G G - - - T C T G G A A G A A G A T G C T G A A 678
820 G C G T T C T T T G A C A T G C C T A C C A A G G C T T T T G C A G T G G T G A T G T G A T A A G G A T G C C T G G 879
679 G G A C T G G C C C T T T C G G G C T A T G C A T A A A G A C T G A T T G T T G C C A G T T C T A C T C T A A A 738
880 G C T G T G C C C A C T T C A T C G A A C A G G G C A T T A A T G T T G T C T C T C C C A A T C A T A T G C C A A G 939
739 A A C T T T G G C C T G T A C A A C A G A G C G T G T G G C G T T G T A C T C T G T T G C T G C C G A C A G T G A A 798
940 A A C A T G G G C T T A T A T G T G A G C G T G A G G A G C C T T C A C T A T A T G T C T G C A A A G A T G C G G A T 999
799 A C G T T T G A T C G C G C A T T C A G C C A A A T G A A A G C G G C G A T T C G C G T A C T A C T C T A A C C C A 858
1000 G A A G C C A A A G G G T A G A G T C A C A G T T G A A G A T C T T G A T C C G T C C C A T G A T T C C A A C C C T 1059
859 C A G A C A C A G C G C G C T T C T G T T T G C C A C C A T C C T G A G A A C A A G A T G C G T T A C G T G C C A T T 918
1060 C C C C T C A A T G G G C C C G G A T T G C T G C T G C C A T T C T G A A C A C C C C A G A T T T G C G - A A A A C A 1118
919 T G G A A C A G A G C T G A C T A T A T G C C C A G C G T A T T C A G C G T A T G C G T A G T T G T T C G T C 978
1119 T G C T G C A A A G T G A A A G C A T G G C G A C C G C A T C A T T G C A T G C G G A C T C A A C T G T C 1178
979 A A T A C G C T G C A G A A A A A G C G A A A C C G C G A C T T C A G C T T T A T C A T C A A A C A G A A C G C G 1038
1179 T C C A A C C T C A A G A A G A G G G T T C C A C C C A C A A C T G G C A A C A C A C A C A C C G A A A T T G G C 1238
1039 A T G T T C T C T T C A G T G G C C T G A C A A A A A A A G A A A G A A G T G C T G C G T T G C C G A A G A G T T T G G C 1098
1239 A T G T T C T G T T C A C A G G C T A A A A G C C T G A A C A G T G G A G C G G T G A T C A A A G G A G T T C T C C 1298
1099 G T A T A T G C G G T T C T T C T G T C G C G T A A A A T G T G C C G G A T G A C A C C A G A T A A C A T G C T 1158
1299 A T C T A C A T G A C A A A G A T G G C C G C A T C T C T G T G C A G G G T C A C C T C C A C A A C A G T G G C 1358
1159 C C G C T G T G C A A G C G A T T 1176
1359 T A C C T T G C C C A T G C C A T T 1376

RESULT 13

CNS0A2EC 1599 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTGH60ZA09 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX828177
VERSION BX828177.1 GI:42459989
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1599)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quefieri,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1599)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.

FEATURES
source
1..1599
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTPGH60A09"
/tissue_type="Hormone Treated Callus"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1599)
/gene="At4g31990"

ORIGIN

Query Match 17.2%; Score 205.4; DB 4; Length 1599;
Best Local Similarity 49.5%; Pred. No. 2.le-52;
Matches 586; Conservative 0; Mismatches 591; Indels 6; Gaps 2;

Qy 4 TTGAGACATTACCGCGCTCTCGCGACCGATTCTGGCGCTGGCGGATCTGTTTGGT 63
Db 218 TTCGAGGTATTAATCTATGCTCCACGACCCCTATCTTGAGTCAGTTAAGCATTCAA 277

Qy 64 GCCGATGAACGTCGCCGCAAAATTAACCTCGGATTTGGTCTATAAAGATGAGACGGGC 123
Db 278 GCTGACACTAAGCGGATGAACCTCAATCTTGGTGTGGTCTTATCGTACTGAGGAATC 337

Qy 124 AAAACCCGGTACTGACGAGGTGAAGAGGTGAACAGTATCTGCTCGAAATGAAC 183
Db 338 CAGCCTTATGTGCTTAATGTTGTTAAAGAGCGGAGAAATTTGATGTTGGAGAGGAGAT 397

Qy 184 ACCAAAATTAACCTCGGCAATGACGGCATCTGTAATTTGGTCTGCTCAGCACTG 243
Db 398 AACAAAGATATCTTCCAAATGAGGGTTGGCAGCATTCACAGGCTACTGCTGAGTTG 457

Qy 244 CTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTGCTCGCAGCGCACAGACTCCG 303
Db 458 CTATTTGGAGCTGGTCACTCTGTTATTAAAGAAACAAAGATAGCAACAATTCAGGGTCTT 517

Qy 304 GGGGCACTGGCGCACTACGGTGGCTGCCGATTTCTCGGCAAAATAATACAGCGTTAAG 363
Db 518 TCGGGAACAGGTTCACTCGGATAGCAGCGCTCTTATAGACGGTATTTCCCTGGAGCA 577

Qy 364 CGTGTGGGTGAGCAACCCAGCTGGCGAACCATAGAGCGTCTTAACCTCGCAGGT 423
Db 578 AAAGTTGTGATCTCATCAACCACTGGGGTAATCACAAGATATCTTCAATGATGCCAAA 637

Qy 424 CTGGAAGTTTCGTGAATACGCTTATTATGATCGGAAATACACACTTTGACTTCGATGCA 483
Db 638 GTTCGGTGGTCCGAATACCGCTACTATGATCCAAACAAATGGTTGGATTTGAGGGA 697

Qy 484 CTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTGAGTGTGTTCCATGGCTGTGC 543
Db 698 ATGATAGCAGATATAAAGGAAGCTCCAGAAGGATCTTCATCTTGTTCACGGATGTCT 757

Qy 544 CATTAACCAACCGGTATCGACCTTACGCTGGAACATGGCAAACTAGTGGCAACTCTCC 603
Db 758 CACAAACCACTGAAATGACCCCAACACAGTGGGTAAATAATGCTGATGTCTATT 817

Qy 604 GTTGAGAAAGGCTGGTTACCGCTGTTTGGCTTACCGAGGGTTTGGCCCGTGG--GT 660

Db 818 CAGGAAAGAACCATATCCCATTTTCGATGTTGCATACACAGGGCTTTGCTAGTGAAGC 877
Qy 661 CTGGAAGAAGATGCTGAGGACTCGCGCTTTTCGGCGCTATGCATAAAGAGCTGATGTT 720
Db 878 CTTGATGAAGATGAGCATCTGTGAGATATTTCCTGAGCGGGGAATGGAGTTTGTGTT 937
Qy 721 GCCAGTTCTCTACTCTAAAAACCTTTGGCCTGTACACAGCAGCGTGTGGCGCTTTGTA 780
Db 938 GCTCAGTCATATAGTAAATAATTTAGTTTGTATGCAGAAAGAAATTTGGGCAATCAATGTC 997
Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGAGCGCGATTCGC 840
Db 998 GTGTGCTCTCAGCTGATGCTGTACAGGGTCAAGAGCCAGTTTGAAAAGGATTTGCTCGG 1057
Qy 841 GCTAACTACTCTAACCCACAGCACAGCGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
Db 1058 CCTATGACTCGAATCCACAGTTTCATGGGGGAGAAATCGTGGCCAAATGTAGTGGGTGAT 1117
Qy 901 GATCGCTTACGTGGCAATTTGGGAACAAGAGCTGATGATATGCGCGCAGCGTATTTCAGGCT 960
Db 1118 GTAATATGTTTCAGTGAATGGAAGCAGAGATGGAATGATGGCAGGAAGAATAAGAGC 1177
Qy 961 ATGCGTCAGTTGTTGTCATACCTGCGAGAAAAGCGCAAAACCG--CGACTTCAGC 1017
Db 1178 GTTAGACAAGAGCTGTATGATAGCTCGTTTCAAAGACAAGAGCGGAGGACTGGTCC 1237
Qy 1018 TTTATCATCAACAGAGCGCATGTTCTCTTCATGTCAGTGCCTGCACAAAAGACAACTGCTG 1077
Db 1238 TTCATTCTGAAGCAAAATGGGCATGTTCTCTTTCACCGGCCCTTAAACAAAGCTCAGAGCGAT 1297
Qy 1078 CGTCTGCGCAAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGCTAAATGTGGCCGG 1137
Db 1298 AACATGAGCAAAATGGCATGTGTATATGACTAAAGCGGAGGATATCATTTGCGCCGA 1357
Qy 1138 ATGACACAGATAACATGGCTCCGCTGTGCGAAGCGATTTGGG 1180
Db 1358 TTATCTCTGGCAAAATCGAGTATCTTGTCTGATCGATCATCG 1400

RESULT 14
CK259742 977 bp mRNA linear EST 03-AUG-2004
LOCUS EST705820 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAB141 5' end, mRNA sequence.
CK259742
ACCESSION CK259742.1 GI:39816720
VERSION CK259742.1
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 977)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: EST705818 EST705819 EST705821
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers
FEATURES
source
1..977
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB141"
/tissue_type="abiotic stress treated leaf and root tissue"


```

/lab host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes=vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

```

ORIGIN

Query Match	17.0%;	Score 203;	DB 7;	Length 977;
Best Local Similarity	51.3%;	Pred. No. 1e-51;		
Matches 498;	Conservative	0;	Mismatches 470;	Indels 3; Gaps 1;
Qy 208	GGCATCCCTGAATTTGGTGCCTGCACTCAGGAACCTGCTTTGGTAAAGTAGGCGCCCTG	267		
Db 2	GACTTGTCTGGAATTTAAACAAATTCAGTGCACCAAGCTTATATTTGGTGTGACAGCCCTGCC	61		
Qy 268	ATCAATGACAACTGTCGACCGCACAGACTCGGGGGGCACTGGCGCACTACGGGTG	327		
Db 62	ATTAAGAGAGATAGGGTGACCACTGTTTCAGTGTCTTGGCGGCACTGGTCTTTGAGGGTT	121		
Qy 328	GCTGCGAATTTCTGGCAAAATAACACCGCTTAAGCGTGTGGGTGAGCAACCAAGC	387		
Db 122	GGTGTGATTTCTGGCAGGCAATATCATGAACATATATATATACACAGCAACA	181		
Qy 388	TGGCGGAACCAATAGAGCGTCTTTAACTCTGAGGTCTGGAAGTTCGTAATAGCTTAT	447		
Db 182	TGGGGAACCACTCTAAGATTTTCACTTTAGCTGGCTTTCAGTAAATATTTATCGTTAC	241		
Qy 448	TATGATCGGAATATCACACTCTTGACTTCGATCGACTGATTAACAGCGCTGAATGAAGCT	507		
Db 242	TATGACCCAGCAACACAGGCGCTAGATTTTCAAGGATTTTGGATGATCTTACGGCTGCA	301		
Qy 508	CAGGCTGGCAGGTAGTCTGTTCATGGCTGCTGCCATAACCAACCGGTATCGACCTT	567		
Db 302	CCCTCTGTGCAATAGTTCTTCTCCATGCTTGTCTCATACCACTGGCGTTGATCCA	361		
Qy 568	ACGCTGGAACCAATGCAAACTAGGCAACACTCTCCGTTGAGAAGGCTGTTACCGCTG	627		
Db 362	ACAATTGACAGTGGGAAAAAATCAGGAAGTTGATGAGGTCCAAGGGGTTGTACCTTTC	421		
Qy 628	TTTGACTTCGCTTACAGGTTTGGCCGTTGG---TCTGGAAGAGATGCTGAAGGACTG	684		
Db 422	TTTGACAGTGTATACAGGTTTGGCCAGTGGAAACCTAGATGACAGTATCTGTT	481		
Qy 685	CGCGTTTTCGGGCTTATGCATAAGAGCTGATTTGTGCGAGTCTCTACTCTAAAACTTT	744		
Db 482	CGCATGTTTGTGGCTGATGGTGGTGAATGTCCTTGACGCTCAGAGTTATGCCAAAACATG	541		
Qy 745	GGCCTGTAAACAGAGCGTGTGGCGCTTGTACTCTGTTGTCGCGACAGTGAAACCGTT	804		
Db 542	GGACTGTATGAGAGCGGTGTCGGTCCCTAGCATTTGTTGCAAAAAATAGCAATGTTGCA	601		
Qy 805	GATCGCGATTCAGGCCAATGAAGCGCGGATTCGCGCTACTCTTAACCCACGCA	864		

Db 602	AGCAAAGTCGAAAGCCAGCTAAAGATAGTTATTAAGGCCAATGTACTCAAAATCCACCGATT	661		
Qy 865	CACGGCGCTCTCTGTTGTTGCCACCATCTCCGAGCAACGATCGGTACGTGCGATTTGGGAA	924		
Db 662	CATGGTGGCTCTATTGTTGCCACTATATCTCAAGGACAGACAAATGTATGACGAATGGACA	721		
Qy 925	CAAGAGCTGACTCATATGCGCCAGCGGTATTCAGCGTATGCGTTCAGTTGTTTCGTCAATACG	984		
Db 722	GTTGAGCTGAAAGCAATGGCCGACCGAATTTATTAGCATGCGCCAGCAACTCTATGATGCC	781		
Qy 985	CTCAGGAAAAAGCGCAAAACCCGACTTTCAGCTTTATATCATCAAAAGACGCGATGTTTC	1044		
Db 782	ATCAAGCTCAAGGTACTCCAGGTGATTGGAGTCATATCATCAAGCAAAATTTGGGATGTTT	841		
Qy 1045	TCCTTATGTCCTTGACAAAGAACAAAGTCTCGGTCTGCGGAGAGTTTGGCGTATAT	1104		
Db 842	ACTTTACAGGATTAAGTTGAGCAAGTCTCGTTCATGACTAGAGAGTATCACATTTAC	901		
Qy 1105	GCGGTTGCTTCTCGTCGCTAAATGTTGGCGGATGACACACAGATACATATGCGTCCGCTG	1164		
Db 902	ATGACATCTGATGGGAGAAATTAGCATGGCTGGTCTTAGTTCTCGGACAGTTTCTCATCTT	961		
Qy 1165	TGCGAAGCGAT 1175			
Db 962	GCTGATGCCAT 972			

RESULT 15
 CENS092WQ 1486 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSLTFR192F03 of Flowers and buds of strain col-0 of Arabidopsis
 thaliana (thale cress).
 ACCESSION BX829537.1 GI:42456349
 VERSION H7C; GSLT cDNA.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1486)
 AUTHORS Castellani, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1486)
 Genoscope.
 DIRECT SUBMISSION
 TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castellani
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 UKV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 Location/Qualifiers
 1. .1486
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"

FEATURES
 source


```

/clone="GSL7FB19ZF03"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..1486
/gene="At5g19550"

```

ORIGIN

```

Query Match      16.8%; Score 200.6; DB 4; Length 1486;
Best Local Similarity 50.9%; Pred. No. 6.6e-51;
Matches 605; Conservative 0; Mismatches 574; Indels 10; Gaps 5;

QY 10 AACATTACCGCGCTCTCCGACCCGATTTCTGGCGTGGCCGATCTGTTCTGCCGAT 69
DB 108 AACGTTGCTCGTGTCTCCGAGATCTATTCTCGGTGTAACCTGTTCTTACAAATGAT 167
QY 70 GAACGTCCTCCGCAAAATTAACCTCGGATGTTGTTCTATAAGATGAGACGGGAAAACC 129
DB 168 CCTAGTCTCTGTTAAGATCAATTTGGGTGTCGTGCTACCGAACTGAGGAAGGAGCCT 227
QY 130 CCGGTACTGACACGCGTGAAGAGCTGAACAGTATCTGCTCGAAAAATGAACCCAC--C 186
DB 228 CTTGTTCTTGATGTGTTGGTGAAGCAGAGCAACAGCTAGTAGAACGATCCGTCTCGGGTC 287
QY 187 AAAAAATTACCTCGGCAATTGACGGCATCCCTGAATTTGGTCTGCTGCACTCAGGAATGCTG 246
DB 288 AAGGAATACATTCCTCATTTGTTGTTATTTCCGATTTTAACAAGTTAAGCGCCAAAGCTCATC 347
QY 247 TTTGTTAAAGGTAGCGGCTCTGATCAATGACAAAGCTGCTCGCAGCGCACAGACTCCGGGG 306
DB 348 TTAGGTGCTGACAGTCTCTCGCATTAAGAAAGTAGAGTTACTACAGTCCAGTCTGCTGTCT 407
QY 307 GGCACCTGGCGCACTACGCTGCTGCTCGATTTCTCGCAAAAAATACACAGGTTAAGCGT 366
DB 408 GGTACTGTTCTTGAGAGTTGGGCTGAGTTTCTCAAAAACACTACACCAAAAGTGTCT 467
QY 367 GTGTGGGTGAGCAACCCAGCTGGCGCAACCAT--AAGAGCGCTTTTAACTCTGCAAGGTCT 425
DB 468 ATTTACATTTCCAAACCAACTTTGGGGAAACCATCCCAAGTTTCAACTTGGCAGGCTT 527
QY 426 GGAAGTTCTGTAATACGCTTATTATGATGCGGAAATACACTCTTGACTTTCGATGCACT 485
DB 528 GTCTGTGGAGTATTTCCGCTACTATGATCTCTCAACCCGCTGCTTGTACTTCAAAAGGCTT 587
QY 486 GATTAACAGCTGAATGAAGCTCAGCTGGCGACCTAGTCTGCTTCATGGCTGCTGCCA 545
DB 588 GCTCAGGATCTTGGCGCTGCAACCATCTGGAGCTATGTCTTACTTTCATGATGTTTCA 647
QY 546 TAACCCAAACCGGTATCGACCTTACGCTGGAAACAATGGCAACAACACTGGCACAACTCTCGGT 605
DB 648 CAATCCCACTGGAGTTGACCCAACTCTGACAGTGGGNAACAGATTCGACAGCTGATGAG 707
QY 606 TGAGAAAGCGTGTGTTACCGCTGTTTGAATTCGTTTACAGGGTTTGGCCGTTGTT--CT 662
DB 708 ATCTAAAGACTTATTACCCCTTTTGTATAGTGCATATCAGGGTTTGTAGTGGTAGCCT 767
QY 663 GGAAGAAGTCTGAGGACTCGCGCTTTCGCGCTATGCATAAAGAGCTGATTTGTTGC 722
DB 768 TAAAAACAGATGCACAATCTGTCCGTTACATTTGTCTGATGGCGGTGAATGCTTGATAGC 827
QY 723 CAGTTCTCTACTTAAAAACTTTGGGCTGTACAAACGAGGCTGTTGGCGCTTGTACTCTGTT 782
DB 828 TCAAGTTTACGCCAAAATATGGGACTTTATGGGAGGCGTGTGTTGGCCCTTAGCATTTGT 887
QY 783 TGCTGCCGACAGTGAACCGTTGATCGCGCAATTCAGCCAAATGAAGCGCGGATTCGCGC 842
DB 888 CTGCAAGTCAGCAGATGTGGCTAGTAAGGTTGAGAGCCAGGTGAAGCTTTGTTGTGCGGCC 947
QY 843 TAACTACTCTAACCCACAGCA--CACGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAACG 901
DB 948 CATGTATTTCGAGCCCCCTTATTCATGAGCATCAATTTGTTGCCACCATCTCTAAAAAGCA 1007
QY 902 ATGCGTTACGTGCGGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTCA--GCG 959

```

Search completed: March 15, 2006, 17:21:02
Job time : 5405 secs

```

DB 1008 GTGATATGTACAACAACTGGACCATCGAGCTGAAGAATAATGGTCCGTATATAAGCGCG 1067
QY 960 TATGCGTCAAGTTGTTTCGTCAATACGCTGCAGGAAAAAGCGCAACCGCACTTCAGCTT 1019
DB 1068 CATGCGCCAAACAGTTTATTGAAGCTATACAAGCTATAGAAGGTACACCTGGTTACTGGAGTCA 1127
QY 1020 TATCATCAAAACAGAACCGCATGTTCTCTTCAAGTGGCTGACAAAAAGAACAAAGTGTGCG 1079
DB 1128 TATTATCAAAACAGATTGGGATGTTTACATTTACTGGATTGAACAGGACAAAGTTGAATT 1187
QY 1080 TCTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTGCGCGGAT 1139
DB 1188 CATGACCAAAAGAGTTCCACATTTACATGACTTCTGATGGGAGAAATAGCATGCGAGGTCT 1247
QY 1140 GACACCAGATAACATGGCTCGCTGTGCGAAGCGATTGTGSCAGTGTCTG 1188
DB 1248 AAGTTGGAAGACAGTGCCTCACCTTGGCGATGCTATGATGCTGCAGTG 1296

```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 13:59:01 ; Search time 6319 Seconds
(without alignments)
10713.808 Million cell updates/sec

Title: US-10-673-786A-1
Perfect score: 1191
Sequence: 1 atgttgagacattaccgc.....cgattggcagtgctgtaa 1191

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_in: *
3: gb_env: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pr: *
9: gb_ro: *
10: gb_sts: *
11: gb_sy: *
12: gb_un: *
13: gb_vi: *
14: gb_htg: *
15: gb_pl: *

ALIGNMENTS

RESULT 1
108485
LOCUS 108485 1293 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 12 from Patent WO 8700202.
ACCESSION 108485
VERSION 108485.1 GI:588805
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Edwards, M.R., Taylor, P.P., Hunter, M.G. and Fotheringham, I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 12 15-JAN-1987;
FEATURES
source
1..1293
/organism="unknown"
/mol_type="unassigned DNA"

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1191	100.0	1293	6	108485 Sequence 12
2	1191	100.0	1331	1	ECASPCG
3	1191	100.0	1331	6	AR095441
4	1191	100.0	1331	6	E16758
5	1191	100.0	1415	1	ECASPC
6	1191	100.0	1468	6	I08484
7	1191	100.0	3659	6	I08487
8	1191	100.0	11470	1	D90731
9	1191	100.0	14759	6	AX370270
10	1191	100.0	20604	1	D90730
11	1191	100.0	110000	1	U00096.09
12	1170.2	98.3	110000	1	AE005174_11
13	1170.2	98.3	110000	1	AE005174_12
14	1170.2	98.3	110000	1	BA000007_11
15	1168.6	98.1	110000	1	AE005674_09
16	1168.6	98.1	292504	1	AE016981
17	1159	97.3	301276	1	AE016758
18	864.6	72.6	110000	1	CP000026_18

Query Match	100.0%;	Score 1191;	DB 6;	Length 1293;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1191;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGTTTGAGACATTACCGCGCTCTGCGACCGATTCTGGCGCTGGCGGATCTGTTT	60	
Db	12	ATGTTTGAGAACATTACCGCGCTCTGCGACCGATTCTGGCGCTGGCGGATCTGTTT	71	
QY	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG	120	
Db	72	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG	131	
QY	121	GGCAAAACCCGGTACTGACCGCTGAAAGGCTGAAACAGTATCTGTCGAAATGAA	180	
Db	132	GGCAAAACCCGGTACTGACCGCTGAAAGGCTGAAACAGTATCTGTCGAAATGAA	191	
QY	181	ACCACAAAATTACCTCGGATTGACCGATCCCTGAATTTGGTGCCTCAGTAA	240	
Db	192	ACCACAAAATTACCTCGGATTGACCGATCCCTGAATTTGGTGCCTCAGTAA	251	
QY	241	CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGACAAACGTCGTCGACGACAGCT	300	

```
|||||
252 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAACGCTGCTCGACGGCAGACT 311
|||||
301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGTGCAAAAAATACCAAGCTT 360
|||||
312 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTGTGCAAAAAATACCAAGCTT 371
|||||
361 AAGCGTGTGTGGTGAGCAACCCCAAGTGGCCGAAACGATGAAGCGTCTTTAACTCTGCA 420
|||||
372 AAGCGTGTGTGGTGAGCAACCCCAAGTGGCCGAAACGATGAAGCGTCTTTAACTCTGCA 431
|||||
421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
|||||
432 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 491
|||||
481 GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCATGCTGC 540
|||||
492 GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCATGCTGC 551
|||||
541 TGCCTAAACCCCAACCGGTATCGACCCCTACGCTGGAACAATGGCAAACTGGGCACAATC 600
|||||
552 TGCCTAAACCCCAACCGGTATCGACCCCTACGCTGGAACAATGGCAAACTGGGCACAATC 611
|||||
601 TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGCCCGTGGT 660
|||||
612 TCCGTTGAGAAAGCGTGGTTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGCCCGTGGT 671
|||||
661 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGCGCTATGCTAAGAGCTGATGTTT 720
|||||
672 CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTCGCGCTATGCTAAGAGCTGATGTTT 731
|||||
721 GCCAGTTCCTACTCTTAAAACTTTGGCGCTGTACAAAGAGCGTGTGGCGTTTGTACTCTG 780
|||||
732 GCCAGTTCCTACTCTTAAAACTTTGGCGCTGTACAAAGAGCGTGTGGCGTTTGTACTCTG 791
|||||
781 GTTGTCTGCCACAGTGAACCGTTGATCGCGCATTCAGCCCAATGAAGCGCGATTTCG 840
|||||
792 GTTGTCTGCCACAGTGAACCGTTGATCGCGCATTCAGCCCAATGAAGCGCGATTTCG 851
|||||
841 GCTAACTACTCTTAACCCACGACACAGCGCGCTTCTGTTGTCGCCACCATCTTGAGCAAC 900
|||||
852 GCTAACTACTCTTAACCCACGACACAGCGCGCTTCTGTTGTCGCCACCATCTTGAGCAAC 911
|||||
901 GATGCGTTACGTGGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
|||||
912 GATGCGTTACGTGGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 971
|||||
961 ATGCGTCAGTTGTTTCGTCATACGCTGCAGGAAAAAGCGCAACCGCGACTTCAGCTTT 1020
|||||
972 ATGCGTCAGTTGTTTCGTCATACGCTGCAGGAAAAAGCGCAACCGCGACTTCAGCTTT 1031
|||||
1021 ATCATCAACAGAACGCGCATGTTCTCCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT 1080
|||||
1032 ATCATCAACAGAACGCGCATGTTCTCCTTCAGTGGCCTGACAAAGAACAAAGTGTGCGT 1091
|||||
1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGGTCGCGTAAATGTGGCGGGATG 1140
|||||
1092 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGGTCGCGTAAATGTGGCGGGATG 1151
|||||
1141 ACACCATCAATGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCCTGATAA 1191
|||||
1152 ACACCATCAATGCTCCGCTGTGCGAAGCGATTGTGGCAGTGCCTGATAA 1202
|||||
```

```
RESULT 2
ECASPCG ECASPCG 1331 bp DNA linear BCT 12-SEP-1993
LOCUS E. coli aspC gene for aspartate aminotransferase.
DEFINITION X05904
ACCESSION X05904
VERSION X05904.1 GI:41012
KEYWORDS aspartate aminotransferase; aspC gene.
SOURCE Escherichia coli
ORGANISM Escherichia coli
```

```
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
AUTHORS Kuramitsu.S., Okuno.S., Ogawa,T., Ogawa,H. and Kagamiyama,H.
TITLE Aspartate aminotransferase of Escherichia coli: nucleotide sequence
of the aspC gene
JOURNAL J. Biochem. 97 (4), 1259-1262 (1985)
PUBMED 3897210
REFERENCE
AUTHORS Kagamiyama,H.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-1988) by Kagamiyama H.
COMMENT Data kindly reviewed (07-MAR-1988) by Kagamiyama H.
FEATURES
source
1..1331
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
10..1200
/note="unnamed protein product; aspartate aminotransferase
(AA 1 - 396)"
/codon_start=1
/transl_table=11
/protein_id="CAA29333.1"
/db_xref="GI:41013"
/db_xref="GOA:P00509"
/db_xref="UniProt/Swiss-Prot:P00509"
/translation="MFENITAAPADPILGLADLFRADERPGKINLIGIYVKDETGTKTP
VLTYSVKAEQYILLENTTKYLDIGIPFGRCTQELLFKGSALINDKRARTLPTFG
GTGALRVADFLAKNTSVKRVVSNPNSHNKSVFNSAGLEVREYAYDAENHTLDFD
ALINSLEAQAQGVVLFHGCCHNPTGIDPTLQWQOTLAQSLKGLPLDFAYQGFPA
RGLEDAELGRAPAAHMKELIVASSYKSNFGLVNERVGACTLVAADSETVDRAFSQMK
AAIRANYSNPPAHGASVVATILSNDALRAIMEQELTDMQRQIRMQLFVNTLQEKGA
NDFDFIIONKWFSGLTKEQVLNREEFGYVAVASRVNVAGTTPDNMAPLCEAI
VAVL"
old_sequence 90
/note="c was a in [1]"
/citation=[1]
old_sequence 502..503
/note="ag was tc in [1]"
/citation=[1]
old_sequence 612
/note="c was g in [1]"
/citation=[1]
old_sequence 657
/note="t was g in [1]"
/citation=[1]
old_sequence 693
/note="g was c in [1]"
/citation=[1]
old_sequence 696
/note="c was t in [1]"
/citation=[1]
old_sequence 699
/note="t was g in [1]"
/citation=[1]
ORIGIN
Query Match 100.0%; Score 1191; DB 1; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTTGCGGCTGGCCGATCTGTTT 60
DB 10 ATGTTTGAGAACATTACCGCGCTCTGCGGACCCGATTTGCGGCTGGCCGATCTGTTT 69
QY 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTCTCTATAAAGATGACG 120
DB 70 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGTTCTCTATAAAGATGACG 129
QY 121 GCAAAACCCCGTACTGACCGCGTGAAGGCTGAACAGTATCTGCTCGAAATGAA 180
DB 130 GCAAAACCCCGTACTGACCGCGTGAAGGCTGAACAGTATCTGCTCGAAATGAA 189
```

181 ACCACCAAAATTAACCTCGCATGACGGCATCCCTGAATTTGGTGCCTGCACCTCAGGAA 240
 190 ACCACCAAAATTAACCTCGCATGACGGCATCCCTGAATTTGGTGCCTGCACCTCAGGAA 249
 241 CTGCTGTTGGTAAAGTAGGCGCTGATCAATGACAAACGTCGTGCGCGGACAGACT 300
 250 CTGCTGTTGGTAAAGTAGGCGCTGATCAATGACAAACGTCGTGCGCGGACAGACT 309
 301 CCGGGGGGCACTGCGGCACTACCGTGGCTGCGGATTTCTGCGCAAAATATACAGGTT 360
 310 CCGGGGGGCACTGCGGCACTACCGTGGCTGCGGATTTCTGCGCAAAATATACAGGTT 369
 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCAATAGAGCGTCTTTAACTCTGCA 420
 370 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCAATAGAGCGTCTTTAACTCTGCA 429
 421 GGTCTGGAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
 430 GGTCTGGAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 489
 481 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 540
 490 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 549
 541 TGGCATAACCCAAACCGGTATCGACCCCTACGCTGGAACCAATGGCAACACACTGGCACAACTC 600
 550 TGGCATAACCCAAACCGGTATCGACCCCTACGCTGGAACCAATGGCAACACACTGGCACAACTC 609
 601 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 660
 610 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 669
 661 CTGGAAGAGATGCTGAAGACTGCGGCTTTTGAATTCGCGCTATGCTAATAAGAGCTGATGTT 720
 670 CTGGAAGAGATGCTGAAGACTGCGGCTTTTGAATTCGCGCTATGCTAATAAGAGCTGATGTT 729
 721 GCCAGTTCCTACTCTAAATACTTTGGCTCTGACAGCGGTGTGGCGCTTGTACTCTG 780
 730 GCCAGTTCCTACTCTAAATACTTTGGCTCTGACAGCGGTGTGGCGCTTGTACTCTG 789
 781 GTTGCTGCCGACAGTGAACCGTTGATTCGCGCAATTCAGCCAAATGAAGCGGCGATTCGC 840
 790 GTTGCTGCCGACAGTGAACCGTTGATTCGCGCAATTCAGCCAAATGAAGCGGCGATTCGC 849
 841 GCTAACTACTCTAAACCAACAGCAGCAGGCGCTTCTGTTGTCGCCACCATCTCTGAGCAAC 900
 850 GCTAACTACTCTAAACCAACAGCAGCAGGCGCTTCTGTTGTCGCCACCATCTCTGAGCAAC 909
 901 GATGCGTTACGTGCGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
 910 GATGCGTTACGTGCGATTTGGGAACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 969
 961 ATGCGTCAGTGTGTCGTCATACGCTGACGAAAAAGGCGCAACCCGCGACTTCAGCTTT 1020
 970 ATGCGTCAGTGTGTCGTCATACGCTGACGAAAAAGGCGCAACCCGCGACTTCAGCTTT 1029
 1021 ATCATCAACAGAAAGCGCATGTTCTCTTCAGTGGCTGACAAAGAAACAAAGTGTGCGT 1080
 1030 ATCATCAACAGAAAGCGCATGTTCTCTTCAGTGGCTGACAAAGAAACAAAGTGTGCGT 1089
 1081 CTGCGGAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGCTTAAATGCGCGGAGT 1140
 1090 CTGCGGAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGCTTAAATGCGCGGAGT 1149
 1141 ACACAGATTAACATGGCTCGCGTGTGCGAAGCGATTTGGCAGTGTGTTAA 1191
 1150 ACACAGATTAACATGGCTCGCGTGTGCGAAGCGATTTGGCAGTGTGTTAA 1200

RESULT 3
 AR095441
 LOCUS

DEFINITION Sequence 23 from patent US 6004773.
 ACCESSION AR095441
 VERSION AR095441.1 GI:10023318
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1331)
 AUTHORS Araki,M., Sugimoto,M., Yoshihara,Y. and Nakamatsu,T.
 TITLE Method for producing L-lysine
 JOURNAL Patent: US 6004773-A 23 21-DEC-1999;
 FEATURES Location/Qualifiers
 1..1331
 /organism="unknown"
 /mol_type="unassigned DNA"

Query Match 100.0%; Score 1191; DB 6; Length 1331;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAAACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGACG 120
 Db 10 ATGTTTGAGAAACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGACG 129
 Qy 61 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGACG 120
 Db 70 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGACG 129
 Qy 121 GGCAAAACCCGGTACTGACAGCGTGAAAGAGCTGAACAGTATCTGCTCGAAAAATGAA 180
 Db 130 GGCAAAACCCGGTACTGACAGCGTGAAAGAGCTGAACAGTATCTGCTCGAAAAATGAA 189
 Qy 181 ACCACCAAAATTAACCTCGGATTTGAGCGCATCCCTGAATTTGGTGCCTGCACCTCAGAA 240
 Db 190 ACCACCAAAATTAACCTCGGATTTGAGCGCATCCCTGAATTTGGTGCCTGCACCTCAGAA 249
 Qy 241 CTGCTGTTGGTAAAGTAGGCGCTGATCAATGACAAACGTCGTGCGCGGACAGACT 300
 Db 250 CTGCTGTTGGTAAAGTAGGCGCTGATCAATGACAAACGTCGTGCGCGGACAGACT 309
 Qy 301 CCGGGGGGCACTGCGGCACTACGCGTGGCTGCGATTTCTGCGCAAAATATACAGGCTT 360
 Db 310 CCGGGGGGCACTGCGGCACTACGCGTGGCTGCGATTTCTGCGCAAAATATACAGGCTT 369
 Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCAATAGAGCGTCTTTAACTCTGCA 420
 Db 370 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCCGAAACCAATAGAGCGTCTTTAACTCTGCA 429
 Qy 421 GGTCTGGAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
 Db 430 GGTCTGGAATTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 489
 Qy 481 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 540
 Db 490 GCACCTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGACGTAGTGTGTTCCATGGCTGC 549
 Qy 541 TGGCATAACCCAAACCGGTATCGACCCCTACGCTGGAACCAATGGCAACACACTGGCACAACTC 600
 Db 550 TGGCATAACCCAAACCGGTATCGACCCCTACGCTGGAACCAATGGCAACACACTGGCACAACTC 609
 Qy 601 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 660
 Db 610 TCGGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCGCTGGT 669
 Qy 661 CTGGAAGAGATGCTGAAGACTGCGGCTTTTGAATTCGCGCTATGCTAATAAGAGCTGATGTT 720
 Db 670 CTGGAAGAGATGCTGAAGACTGCGGCTTTTGAATTCGCGCTATGCTAATAAGAGCTGATGTT 729
 Qy 721 GCCAGTTCCTACTCTAAATACTTTGGCTCTGACAGCGGTGTGGCGCTTGTACTCTG 780
 Db 730 GCCAGTTCCTACTCTAAATACTTTGGCTCTGACAGCGGTGTGGCGCTTGTACTCTG 789

QY 781 GTTGCTGCCAGCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCCG 840
Db |||||
QY 790 GTTGCTGCCAGCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCCG 849
Db |||||
QY 841 GCTAACTACTCTAACCCACGACACAGCGCGTCTCTGTTGTTGCCACCATCTCTGAGCAAC 900
Db |||||
QY 850 GCTAACTACTCTAACCCACGACACAGCGCGTCTCTGTTGTTGCCACCATCTCTGAGCAAC 909
QY 901 GATCGCTTACGTCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 960
Db |||||
QY 910 GATCGCTTACGTCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 969
QY 961 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAGGCGCAAAACCGCGACTTTCAGCTTT 1020
Db |||||
QY 970 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAGGCGCAAAACCGCGACTTTCAGCTTT 1029
QY 1021 ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGCCTGACAAAAGAACAAAGTGTGCGT 1080
Db |||||
QY 1030 ATCATCAAAACAGAACGGCATGTTCTCTTCAGTGGCCTGACAAAAGAACAAAGTGTGCGT 1089
QY 1081 CTCGCGGAAGAGTTTGGCGTATATGCGTTGCTTCTGTCGCGTAAATGTGGCGCGGATG 1140
Db |||||
QY 1090 CTCGCGGAAGAGTTTGGCGTATATGCGTTGCTTCTGTCGCGTAAATGTGGCGCGGATG 1149
QY 1141 ACACAGATAACATGCTCGCTGTCGGAAGGATTTGTGCGAGTGTCTGTA 1191
Db |||||
QY 1150 ACACAGATAACATGCTCGCTGTCGGAAGGATTTGTGCGAGTGTCTGTA 1200

RESULT 4

E16758 1331 bp DNA linear PAT 28-JUL-1999
LOCUS gDNA encoding aspartate aminotransferase (AAT).
DEFINITION E16758
ACCESSION E16758
VERSION 1 (bases 1 to 1331)
KEYWORDS Araki,M., Sugimoto,M., Yoshihara,Y. and Nakamatsu,W.
SOURCE Escherichia coli
ORGANISM Escherichia coli
REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
AUTHORS 1 (bases 1 to 1331)
TITLES Araki,M., Sugimoto,M., Yoshihara,Y. and Nakamatsu,W.
JOURNAL PRODUCTION OF L-LYSINE
COMMENT Patent: JP 1998215883-A 16 18-AUG-1998;
AJINOMOTO CO INC
OS Escherichia coli
PN JP 1998215883-A/16
PD 18-AUG-1998
PF 03-DEC-1997 JP 1997333238
PR 05-DEC-1996 JP 96P 325659
PI ARAKI MASAYUKI, SUGIMOTO MASAKAZU, YOSHIHARA YASUHIKO, PI
NAKAMATSU WATARU
PC C12N15/09,C12N1/21,C12P13/08,(C12N15/09,C12R1:15),(C12N1/21,
C12R1:15),
PC (C12P13/08,C12R1:15);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1331
FT FT /organism='Escherichia coli'
FT FT /strain='JM109'
FT 5'UTR 1..9
FT CDS 10..1197
FT FT /product='aspartate aminotransferase' FT
FT /gene='aspC'
FT 3'UTR 1198..1331.
FEATURES
source
1..1331
Location/Qualifiers
/organism="Escherichia coli"
/mol_type="genomic DNA"

ORIGIN /db_xref="taxon:562"
Query Match 100.0%; Score 1191; DB 6; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGGAGAACATTACCGCGCTCTCCGACCCGATTTCTGGGCTGGCCGATCTGTTT 60
Db |||||
QY 10 ATGTTTGGAGAACATTACCGCGCTCTCCGACCCGATTTCTGGGCTGGCCGATCTGTTT 69
Db |||||
QY 61 CGTCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db |||||
QY 70 CGTCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 129
QY 121 GGCACCAACCCCGGTACTGACGAGCTGAAAAGGCTGAACAGTATCTGCTCAAAATGA 180
Db |||||
QY 130 GGCACCAACCCCGGTACTGACGAGCTGAAAAGGCTGAACAGTATCTGCTCAAAATGA 189
QY 181 ACCACCAAAATTAACCTCGGCAATTCACGCGCATCCCTGAATTTGGTCTGCACTCAGGAA 240
Db |||||
QY 190 ACCACCAAAATTAACCTCGGCAATTCACGCGCATCCCTGAATTTGGTCTGCACTCAGGAA 249
QY 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAACGTCCTCGCACGACACACT 300
Db |||||
QY 250 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAACGTCCTCGCACGACACACT 309
QY 301 CCGGGGGGCACTGGCGCACTAGCGTGGCTGCGGATTTCTGGCAAAAATACCAGCTT 360
Db |||||
QY 310 CCGGGGGGCACTGGCGCACTAGCGTGGCTGCGGATTTCTGGCAAAAATACCAGCTT 369
QY 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGCAACCAATGAAGAGCTCTTTAACTCTGCA 420
Db |||||
QY 370 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCCGCAACCAATGAAGAGCTCTTTAACTCTGCA 429
QY 421 GGTCTGGAAGTTGCTGAATACGCTTTATATGATGCGGAAATACACACTCTTGACTTCGAT 480
Db |||||
QY 430 GGTCTGGAAGTTGCTGAATACGCTTTATATGATGCGGAAATACACACTCTTGACTTCGAT 489
QY 481 GCATGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGTAGTGTCTTCATGCTGCTC 540
Db |||||
QY 490 GCATGATTAACAGCCTGAATGAAGCTCAGGCTGGCGAGTAGTGTCTTCATGCTGCTC 549
QY 541 TGCCATAACCCCAACCGGTATCGACCTAGCTGGAACAATGCGCAACACTTGGCAACACTC 600
Db |||||
QY 550 TGCCATAACCCCAACCGGTATCGACCTAGCTGGAACAATGCGCAACACTTGGCAACACTC 609
QY 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTTCGCCGTGT 660
Db |||||
QY 610 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTTCGCCGTGT 669
QY 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGGGCTATGCATAAAGAGCTGATTGTT 720
Db |||||
QY 670 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTCCGGGCTATGCATAAAGAGCTGATTGTT 729
QY 721 GCCAGTTCTACTCTAAAACCTTTGCGCTGTACACAGCGGTGTTGGCGCTTGTACTCTG 780
Db |||||
QY 730 GCCAGTTCTACTCTAAAACCTTTGCGCTGTACACAGCGGTGTTGGCGCTTGTACTCTG 789
QY 781 GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCCG 840
Db |||||
QY 790 GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCCG 849
QY 841 GCTAACTACTCTAACCCACGACACAGCGCGTCTCTGTTGTTGCCACCATCTCTGAGCAAC 900
Db |||||
QY 850 GCTAACTACTCTAACCCACGACACAGCGCGTCTCTGTTGTTGCCACCATCTCTGAGCAAC 909
QY 901 GATCGTTACGTCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 960
Db |||||
QY 910 GATCGTTACGTCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 969
QY 961 ATGCGCTCAGTTTCTGCTCAATACGCTGCAGGAAAGGCGCAAAACCGCGACTTTCAGCTTT 1020
Db |||||

Db 970 ATCGCTCAGTTGTTGCTCAATACGCTGACGAGAAAAAGGCGCAAAACCGGACTTCAGCTTT 1029
 QY 1021 ATCATCAAAACAGAACGGCATGTTCTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCGT 1080
 Db 1030 ATCATCAACAGAACGGCATGTTCTCTCTTCAGTGGCTGACAAAAGAAACAAGTGTGCGT 1089
 QY 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTGCTTCTCTGCGCTGCGTAAATGTTGGCCGGATG 1140
 Db 1090 CTGCGGAAGAGTTTGGCGTATATGCGGTGCTTCTCTGCGCTGCGTAAATGTTGGCCGGATG 1149
 QY 1141 ACACCAAGATACATGCTCGCTGTCGAGAGGATTTGTCGAGTGTCTGTA 1191
 Db 1150 ACACCAAGATACATGCTCGCTGTCGAGAGGATTTGTCGAGTGTCTGTA 1200

RESULT 5
 ECASPC 1415 bp DNA linear BCT 12-SEP-1993
 LOCUS E. coli aspC gene for aspartate aminotransferase.
 DEFINITION X03629
 ACCESSION X03629.1 GI:41010
 VERSION
 KEYWORDS aminotransferase.
 SOURCE Escherichia coli
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 Fotheringham,I.G., Dacey,S.A., Taylor,P.P., Smith,T.J.,
 Hunter,M.G., Finlay,M.E., Primrose,S.B., Parker,D.M. and
 Edwards,R.M.
 The cloning and sequence analysis of the aspC and tyrB genes from
 Escherichia coli K12. Comparison of the primary structures of the
 aspartate aminotransferase and aromatic aminotransferase of E. coli
 with those of the pig aspartate aminotransferase isoenzymes
 Biochem. J. 234 (3), 593-604 (1986)

JOURNAL PUBMED 3521591
 COMMENT Data kindly reviewed (25-JUN-1986) by M. Hunter.
 FEATURES
 Location/Qualifiers
 1..1415
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /db_xref="taxon:562"
 68..73
 /note="pot. promoter signal"
 91..96
 /note="pot. promoter signal"
 104
 /note="mRNA start site"
 127..130
 /note="ribosome binding site"
 138..1328
 /note="unnamed protein product; aspartate aminotransferase (aspC)"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA27279.1"
 /db_xref="GI:41011"
 /db_xref="Gene: P00509"
 /db_xref="UniProt/Swiss-Prot: P00509"
 /translation="MFENITAAPADPILGLADLFRADERPGKINILGIVYKDETKTP
 VLTYSKASQYLENETTKNYLIDIGIPFGRCTOELLFPGKSALINDKRARTQTPG
 GTGALRVADFLAKNTSVKRVMSNPSPNHSVNSAGLEVREYAYDAENHTLDFD
 ALINLSNEAQAGDVLPFGCCNHPGIDITLQWQTLAQLSVEKWLPLDFAYQGF
 RLGEEDAEGLRAFAAMHKLIVLASSYSKFNGLYNERVAGACTLVAADSETVDRAFQMK
 AAIRANYSPPPHAGASVATILSNDLRAIWEQLTDMRQIRMRQLFVNLTQSKGA
 NRDFSGFIKQNGMFSFSLTKBQLRLREFFGYVAVASGRVNVAGMTPDNMAPLCEAI
 VAVL"

ORIGIN
 Query Match 100.0%; Score 1191; DB 1; Length 1415;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAAACATTACCGCGCTCTCGCGACCCCGATTTCTGGGCGCTGGCGGATCTGTTT 60
 Db 138 ATGTTTGAGAAACATTACCGCGCTCTCGCGACCCCGATTTCTGGGCGCTGGCGGATCTGTTT 197
 QY 61 CGTGGCGATGAACGCTCCCGCAAAATTAACCTCGGATTTGTTCTATATAAGATGAGACG 120
 Db 198 CGTGGCGATGAACGCTCCCGCAAAATTAACCTCGGATTTGTTCTATATAAGATGAGACG 257
 QY 121 GGCAGAAACCCCGGTACTGACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGA 180
 Db 258 GGCAGAAACCCCGGTACTGACACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGA 317
 QY 181 ACACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTGCCTGCACTCAGGAA 240
 Db 318 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTGCCTGCACTCAGGAA 377
 QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGACACAGACT 300
 Db 378 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGACACAGACT 437
 QY 301 CCGGGGGCAGCTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACACAGCTT 360
 Db 438 CCGGGGGCAGCTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACACAGCTT 497
 QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 420
 Db 498 AAGCGTGTGGGTGAGCAACCAAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 557
 QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAAAATACACTCTTGACTTCGAT 480
 Db 558 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAAAATACACTCTTGACTTCGAT 617
 QY 481 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 540
 Db 618 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTTCCATGGCTGC 677
 QY 541 TGCCATAACCCAAACCGGTATCGACCTTACGCTGGAACCAATGGAACACACTGGCACAATC 600
 Db 678 TGCCATAACCCAAACCGGTATCGACCTTACGCTGGAACCAATGGAACACACTGGCACAATC 737
 QY 601 TCGGTTGAGAAAGGCTGGTTACGGCTGTTTGACTTACCGCTTACCGAGGTTTGGCCGCTGGT 660
 Db 738 TCGGTTGAGAAAGGCTGGTTACGGCTGTTTGACTTACCGCTTACCGAGGTTTGGCCGCTGGT 797
 QY 661 CTGGAAGAGATGCTGAAGGACTTGGCGGCTTTCGCGGCTATGCTATGAAGAGCTGATTGTT 720
 Db 798 CTGGAAGAGATGCTGAAGGACTTGGCGGCTTTCGCGGCTATGCTATGAAGAGCTGATTGTT 857
 QY 721 GCCAGTTCTCTCTAAACCTTTGGCCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
 Db 858 GCCAGTTCTCTCTAAACCTTTGGCCTGTACACGAGCGTGTGGCGCTTGTACTCTG 917
 QY 781 GTTGTGCGGACAGTGAACCGTTGATCGCGATTCAGGCAATGAAGCGGCGATTCGC 840
 Db 918 GTTGTGCGGACAGTGAACCGTTGATCGCGATTCAGGCAATGAAGCGGCGATTCGC 977
 QY 841 GCTAACTACTTAACCCACGACACGCGCTTCTGTTGTTGCGCACCATCTCTGAGCAAC 900
 Db 978 GCTAACTACTTAACCCACGACACGCGCTTCTGTTGTTGCGCACCATCTCTGAGCAAC 1037
 QY 901 GATGCGTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGGTATTACGCGT 960
 Db 1038 GATGCGTTACGTGCGATTTGGGAAACAAGAGCTGACTGATATGCGCCAGCGGTATTACGCGT 1097
 QY 961 ATGCGTCAAGTGTGCTCAATACGCTGAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1020
 Db 1098 ATGCGTCAAGTGTGCTCAATACGCTGAGGAAAAAGGCGCAAAACCGCGACTTCAGCTTT 1157
 QY 1021 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTTCACAAAAAGAAACAAGTGTGCGCT 1080
 Db 1158 ATCATCAACAGAACGGCATGTTCTCTTCAGTGGCTTCACAAAAAGAAACAAGTGTGCGCT 1217
 QY 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTCTGCTGCGTAAATGTGCGCGGATG 1140

```
Db 1218 CTCGCGAAGAGTTTGGCGTATATGCGTTGCTTCTGTCGCTAAATGTGGCCGGATG 1277
Qy 1141 ACACAGATAACATGGCTCGCTGTCGGAAGCAATTTGTGGCAGTGTCTGTAA 1191
Db 1278 ACACAGATAACATGGCTCGCTGTCGGAAGCAATTTGTGGCAGTGTCTGTAA 1328

RESULT 6
LOCUS 108484 1468 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 9 from Patent WO 8700202.
ACCESSION 108484
VERSION 108484.1 GI:588804
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1468)
AUTHORS Edwards,M.R., Taylor,P.P., Hunter,M.G. and Fotheringham,I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 9 15-JAN-1987;
FEATURES
    Location/Qualifiers
        source
            1..1468
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1191; DB 6; Length 1468;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTTGAGAACATTAACCGCGCTCTCTGCGGACCGCAATCTTGGCGCTGGCGGATCTGTTT 60
Db 191 ATGTTTGAGAACATTAACCGCGCTCTCTGCGGACCGCAATCTTGGCGCTGGCGGATCTGTTT 250
Qy 61 CGTGCCGATGAACGTCCTCGGCAAAATTAACCTCGGAGTTGGTCTCTATAAGATGAGCG 120
Db 251 CGTGCCGATGAACGTCCTCGGCAAAATTAACCTCGGAGTTGGTCTCTATAAGATGAGCG 310
Qy 121 GGCAAAACCCCGTACTGACAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 180
Db 311 GGCAAAACCCCGTACTGACAGCGTGAAGAGGCTGAACAGTATCTGCTCGAAATGAA 370
Qy 181 ACCACAAATAATTAACCTGGCATTGACCGCATCCCTGAAATTTGGTCTGCTGCACTCAGAA 240
Db 371 ACCACAAATAATTAACCTGGCATTGACCGCATCCCTGAAATTTGGTCTGCTGCACTCAGAA 430
Qy 241 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTGCTCGACGGCACAGCT 300
Db 431 CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGTGCTCGACGGCACAGCT 490
Qy 301 CGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAATAACACGCTT 360
Db 491 CGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAATAACACGCTT 550
Qy 361 AAGCGTGTGGTGAGCAACCCCAAGTGGCCGAAACATAGAGCGTCTTTAACTCTGCA 420
Db 551 AAGCGTGTGGTGAGCAACCCCAAGTGGCCGAAACATAGAGCGTCTTTAACTCTGCA 610
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACAACCTCTGACTTCGAT 480
Db 611 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACAACCTCTGACTTCGAT 670
Qy 481 GCACTGATTAAACAGCCTGAATGAGCTCAGGCTGGGCAAGTAGTGTGTTTCAATGCTGC 540
Db 671 GCACTGATTAAACAGCCTGAATGAGCTCAGGCTGGGCAAGTAGTGTGTTTCAATGCTGC 730
Qy 541 TGCATAACCCCAACCGGATTCGACCTCAAGTGGCAACATGCGCAACATCTGGCAACATC 600
Db 731 TGCATAACCCCAACCGGATTCGACCTCAAGTGGCAACATGCGCAACATCTGGCAACATC 790
Qy 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTACCGGTTTACCGGTTTTCGCCGTGGT 660
```

```
Db 791 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTCGCTTACAGGGTTTTCGCCGTGT 850
Qy 661 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTCGGGGCTATGCATATAAGAGCTGATTGT 720
Db 851 CTGGAAGAAGATGCTGAAGGACTGGCGCTTTCGGGGCTATGCATATAAGAGCTGATTGT 910
Qy 721 GCCAGTTCCTACTCTAAATAAATTTTGGCCTGTACAACAGAGCGTGTGGCGCTTGTACTCTG 780
Db 911 GCCAGTTCCTACTCTAAATAAATTTTGGCCTGTACAACAGAGCGTGTGGCGCTTGTACTCTG 970
Qy 781 GTTGTCTGCCGACAGTGAACCCGTTGATCGCGCAATTCAGCCAAATGAAGCGCGATTCGC 840
Db 971 GTTGTCTGCCGACAGTGAACCCGTTGATCGCGCAATTCAGCCAAATGAAGCGCGATTCGC 1030
Qy 841 GCTAACTACTTAACCCACAGCAGCACCGCGCTTCTGTTTGGCCACCATCTCTGAGCAAC 900
Db 1031 GCTAACTACTTAACCCACAGCAGCACCGCGCTTCTGTTTGGCCACCATCTCTGAGCAAC 1090
Qy 901 GATCGTTACGTCGCAATTTGGGAAACAAGAGCTGACTGATATCGCCAGCGTATTCAGCGT 960
Db 1091 GATCGTTACGTCGCAATTTGGGAAACAAGAGCTGACTGATATCGCCAGCGTATTCAGCGT 1150
Qy 961 ATGCGTCAAGTTGTTGCTCAATACGCTGCAGGAAAAAGCGCAAAACCGCACTTCAGCTTT 1020
Db 1151 ATGCGTCAAGTTGTTGCTCAATACGCTGCAGGAAAAAGCGCAAAACCGCACTTCAGCTTT 1210
Qy 1021 ATCATCAAAACAGAACGGCATGTTCTTCACTGAGCTGCTGACAAAAGAACAGTGTGCGT 1080
Db 1211 ATCATCAAAACAGAACGGCATGTTCTTCACTGAGCTGCTGACAAAAGAACAGTGTGCGT 1270
Qy 1081 CTGCGGAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTAAATGTCGCCGGGATG 1140
Db 1271 CTGCGGAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCTAAATGTCGCCGGGATG 1330
Qy 1141 ACACAGATAACATGCTCCGCTGTCGAAAGCGATTGTGGCAGTGTCTGTAA 1191
Db 1331 ACACAGATAACATGCTCCGCTGTCGAAAGCGATTGTGGCAGTGTCTGTAA 1381

RESULT 7
LOCUS 108487 3659 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 8700202.
ACCESSION 108487
VERSION 108487.1 GI:588807
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3659)
AUTHORS Edwards,M.R., Taylor,P.P., Hunter,M.G. and Fotheringham,I.G.
TITLE COMPOSITE PLASMIDS FOR AMINO ACID SYNTHESIS
JOURNAL Patent: WO 8700202-A 15 15-JAN-1987;
FEATURES
    Location/Qualifiers
        source
            1..3659
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1191; DB 6; Length 3659;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTTGAGAACATTAACCGCGCTCTCTGCGGACCGCAATCTTGGCGCTGGCGGATCTGTTT 60
Db 2344 ATGTTTGAGAACATTAACCGCGCTCTCTGCGGACCGCAATCTTGGCGCTGGCGGATCTGTTT 2403
Qy 61 CTGCGGATGAACGTCCTGGGCAAAATTAACCTCGGAGTTGGTGTCTATAAGATGAGCG 120
Db 2404 CTGCGGATGAACGTCCTGGGCAAAATTAACCTCGGAGTTGGTGTCTATAAGATGAGCG 2463
Qy 121 GSCAAAACCCCGCTACTGACAGCGTGAAGAAAGGCTGAACAGTATCTGCTCGAAATGAA 180
```


Db 2464 GGCAAAACCCGGTACTGACAGCGTGAAAGAGCGTGAAACAGTATCTGCTCGAAATGAA 2523
 QY 181 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGGAATTTGCTGCTGCCTCAGGAA 240
 Db 2524 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGGAATTTGCTGCTGCCTCAGGAA 2583
 QY 241 CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACGTCCTCGCAGGACACAGACT 300
 Db 2584 CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACGTCCTCGCAGGACACAGACT 2643
 QY 301 CCGGGGGCACTCGGCACACTACGGTGCTGCGCATTTCTTGGCAAAAATACACAGGTT 360
 Db 2644 CCGGGGGCACTCGGCACACTACGGTGCTGCGCATTTCTTGGCAAAAATACACAGGTT 2703
 QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 420
 Db 2704 AAGCGTGTGGGTGAGCAACCAAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 2763
 QY 421 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
 Db 2764 GGTCTGGAAGTTCTGTAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 2823
 QY 481 GCACTGATTAAAGCTGAATGAAGCTCAGCTGGCGACGATGCTGTGTTCCATGCTGTC 540
 Db 2824 GCACTGATTAAAGCTGAATGAAGCTCAGCTGGCGACGATGCTGTGTTCCATGCTGTC 2883
 QY 541 TGCCTAATCCCAACCGGTATCGACCTACGCTCGGAACAATGCAACACCTGGGCACAACTC 600
 Db 2884 TGCCTAATCCCAACCGGTATCGACCTACGCTCGGAACAATGCAACACCTGGGCACAACTC 2943
 QY 601 TCGTGTGAGAAAGCGTGTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGGCCCGTGGT 660
 Db 2944 TCGTGTGAGAAAGCGTGTACCGCTGTTTGACTTCGCTTACCAGGCTTTTGGCCCGTGGT 3003
 QY 661 CTGGAAGAAGTCTGAAGACTCGCGCTTTTCGCGCTATGCTAAGAGCTGATGTT 720
 Db 3004 CTGGAAGAAGTCTGAAGACTCGCGCTTTTCGCGCTATGCTAAGAGCTGATGTT 3063
 QY 721 GCCAGTTCCTACTCTAANAATTTGGCCTGTACAAGAGGCTGTGGCGCTGTACTCTG 780
 Db 3064 GCCAGTTCCTACTCTAANAATTTGGCCTGTACAAGAGGCTGTGGCGCTGTACTCTG 3123
 QY 781 GTTGCTGCGCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTGCG 840
 Db 3124 GTTGCTGCGCACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTGCG 3183
 QY 841 GCTAATCTCTTAACCCACAGCACACGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
 Db 3184 GCTAATCTCTTAACCCACAGCACACGGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 3243
 QY 901 GATCGTTACGTGCGATTTGGGACAGAGCTGACTGATATGCGCAGCGTATTACGCGT 960
 Db 3244 GATCGTTACGTGCGATTTGGGACAGAGCTGACTGATATGCGCAGCGTATTACGCGT 3303
 QY 961 ATGCGTCAGTTGTTCTGTAATACGCTGAGGAAAAAGGCGCAACCGCACTTCAGCTTT 1020
 Db 3304 ATGCGTCAGTTGTTCTGTAATACGCTGAGGAAAAAGGCGCAACCGCACTTCAGCTTT 3363
 QY 1021 ATCATCAACAGAACCGCATGTTCTCTTTCAGTGGCTTGACAAAGAACAAAGTGTGCGT 1080
 Db 3364 ATCATCAACAGAACCGCATGTTCTCTTTCAGTGGCTTGACAAAGAACAAAGTGTGCGT 3423
 QY 1081 CTGCGGGAAGAGTTTGGCGTATATGCGGTGCTTCTGTTGCGGTGCGTAAATGTGGCCGGATG 1140
 Db 3424 CTGCGGGAAGAGTTTGGCGTATATGCGGTGCTTCTGTTGCGGTGCGTAAATGTGGCCGGATG 3483
 QY 1141 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGCGAGTGTGTA 1191
 Db 3484 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGCGAGTGTGTA 3534

D90731/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

D90731
 Escherichia coli K12 genomic DNA. (21.2 - 21.5 min). linear BCT 25-DEC-2002
 D90731 AB001340
 D90731.1 GI:1651452
 Complete and shotgun sequencing; ycbB; aspC; ompF; tolF; cmlB; coa; cry; asnS; tss; pncB; pepN; ycbE; yzeB.
 Escherichia coli K12
 Escherichia coli K12
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
 1
 Oshima,T., Aiba,H., Baba,T., Fujita,K., Hayashi,K., Hayashi,K., Honjo,A., Ikemoto,K., Inada,T., Itoh,T., Kajihara,M., Kanai,K., Kashimoto,K., Kimura,S., Kitagawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishio,H., Nishio,Y., Saito,N., Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y., Yano,M.
 A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map
 DNA Res. 3 (3), 137-155 (1996)
 8905232
 2
 Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishio,H., Nishio,Y., Oshima,T., Saito,N., Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.
 The systematic sequencing of the Escherichia coli genome in Japan
 Unpublished
 3 (bases 1 to 11470)
 Mori,H.
 Direct Submission
 Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan
 (E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)
 Collaboration Information:
 Project:
 The Japan E.coli genome DNA sequencing project
 Group:
 The Japan E.coli genome DNA sequencing group
 Members: (1995.4 - 1996.3)
 Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikenoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishio,H., Nishio,Y., Oshima,T., Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.
 Headed by:
 Name: Takashi Horiuchi
 Address: National Institute of Basic Biology, Okazaki, 444, Japan
 E-mail: kishori@nibb.ac.jp
 Information Operator:
 Name: Hirotsada Mori
 Address: NARA Institute of Science and Technology, Ikoma, 630-01, Japan
 E-mail: hmori@gtc.aist-nara.ac.jp
 URL:
 The Japan E. coli genome database
 http://bsw3.aist-nara.ac.jp.
 Location/Qualifiers
 1. .11470
 /organism="Escherichia coli K12"
 /mol_type="genomic DNA"
 /strain="K12"
 /db_xref="taxon:83333"
 /map="21.2-21.5 min"

FEATURES
 source

```
/clone="Kohara clone #219"
/note="Nucleotide position 983395-994864 from the
initiation site of ThrA (0 min.).
This clone is from Kohara lambda miniset library."
1. .47
/gene="ycbb"
CDS
41. .47
/note="ycbb"
/note="ORF ID:0218#11
similar to PIR Accession Number PH0205"
/codon_start=3
/transl_table=11
/product="Hypothetical protein 63 (MukB 3' region)"
/protein_id="BAA35677.1"
/db_xref="GI:4062495"
/translation="SSQIVSKAEQLIR"
228. .776
/note="ORF ID:0219#1
similar to PIR Accession Number F64039"
/codon_start=1
/transl_table=11
/product="Hypothetical protein H11666"
/protein_id="BAA35678.1"
/db_xref="GI:4062496"
/translation="MDKPDNRKLLALGGVALGHAILLTPAFATLSTPRRILLTN
LHTGSIKAEFFDGRGYIOELAKUNHFRDPRANKIKSIDPGLFDQLRLQGLGTR
KEVQLISYRSIDTNELRARGVAKKSYHTKGQMDPHIEGIALSNIRKAALSNRA
CGVGYYPRNSFVHIIDTPARHW"
803. .1450
/note="ORF ID:0219#2
similar to SwissProt Accession Number P54501"
/codon_start=1
/transl_table=11
/product="Hypothetical 23.2 kd protein in soda-comGA
intergenic region."
/protein_id="BAA35679.1"
/db_xref="GI:4062497"
/translation="MNYRIPIVAFSQNSCLIWCEOTRLAALVDPGGDAEKIQEVD
SGLTMLQLLTHDVGAAELAQHGVVFGPEKEDEFWLQGLPAQSRMFLGEEC
QPLTDMRLNEGTDSIGNVTQLVHQPGHTFHVVFDDRAKLISGDIKFGVGVR
SDPFRGDHNLISSIKDLLPLGDDVIFIPGHGPLSTLGYERLHNPFLQDEMPVW"
complement (1672. .2862)
/gene="aspC"
complement (1672. .2862)
/note="aspC"
/note="ORF ID:0219#3
similar to PIR Accession Number A00598"
/codon_start=1
/transl_table=11
/product="Aspartate transaminase (EC 2.6.1.1)"
/protein_id="BAA35680.1"
/db_xref="GI:1651453"
/translation="MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDETGT
VLTSVKKAEQYLEENITKNYLIGDIGIPFGCTOELLFGKGSALINDKRAKTAOTPG
GTALRVAADFLAKNTSVKRVVNSFNHKSQVENSAGLEVREYAYDAENHILDFD
ALINSNEAAGVDFLPHCGCNHPTGIDPTLQWQTLAQLSVEKGLWLPFLDFAYQGF
RLGEEDEAGLRFAPAAHKEILIVASSYKFLGYNRVGACTLVAADSETVDRAFQMK
ALIRANYSNPPAHGASVATILSNALRAIWEOLTDMRQIRQRMQLFVNTLQEKGA
NRDFSFIKQNGMFSGLTKEQVLRLEEFYVYVAGSVNVAGMTPDNMAPLCEAI
VAVL"
complement (3047. .4135)
/gene="cmIB"
/note="synonyms: coa, cry, ompF, tolF"
complement (3047. .4135)
/gene="cmIB"
/note="ORF ID:0219#4
similar to SwissProt Accession Number P02931"
/codon_start=1
/transl_table=11
/product="Outer membrane protein f precursor (outer
membrane protein Ia, Ia, or b)."
/protein_id="BAA35681.1"
/db_xref="GI:1651454"
```

```
gene
CDS
/note="synonym: tsb"
complement (4738. .6138)
/gene="asnS"
/note="ORF ID:0219#5
similar to SwissProt Accession Number P17242"
/codon_start=1
/transl_table=11
/product="Asparaginyl-tRNA synthetase (EC 6.1.1.22)
(asparagine-tRNA ligase) (asnRS)."
/protein_id="BAA35682.1"
/db_xref="GI:1651455"
/translation="MSVVPVADVLOQRVAVDSEVTVRGWVTRDRSKAGISFLAVYDG
SCFDPQAVINNSLPYNEDVLRLTTGCSVITGKVVASPGGQQQFEIOASKEVAGW
VEDPDTYPMAAKRHSIEYLREVAHLRPTNLIGVARVHRHTLAQALHFFNKGQFVW
STPLTASDTGTGAGEMPRVSTLDLENLPRNDQKVDFOKDFKESFLTVSQLNGET
YACALSKIYTFGPTPRAENSTSRHLAEFWMLPEVAFANLNDIAGLAEMLKYPFKA
VLEERADMKFASVDKDAVSRLEFFIEADPAQVDYTDATVILENCGRKAKFNPVWG
VLSSEHRYLAEHFHFKAPVVVKYPKDKIKARYMLNEDGKTVAAMDVLAPGIGETIG
QSREERLDLAEHMLELNKEDYMWYRDRLRYGTVPHSFGFLGFERLIAVYTGQVN
VRDVIFPPTPRNASF"
complement (6307. .7509)
/gene="pncB"
complement (6307. .7509)
/gene="pncB"
/note="ORF ID:0220#1
similar to PIR Accession Number JQ0756"
/codon_start=1
/transl_table=11
/product="Nicotinate phosphoribosyltransferase (EC
2.4.2.11)"
/protein_id="BAA35683.1"
/db_xref="GI:1651456"
/translation="MTQFASPVVLSHLLDTDAYKLMQQAQVPHHYDVHVAEFCRGD
DLGIYADAIQEOVQAOHLQDDYOWLSALPFFKADYLNWLRFRNFPEQVTVSN
DNGKLDIRLSGPMREVILWEVELLAVISPMVHRYSPOADVAQALDTLESKLVDFSA
TAGLMSRFLHMDFGTRRRPSREVETIVKRLQBSWFVGTSNYDLARLSLTDMGTQ
AHEWFOAHQIQSPDLANSORAAALAEELFPQGLALTDCTIMDLFLDFGVEFASR
YQGLRSDSGDPVEWGEKAIHYEKLIGDPQSKTLVPSDNLRLKRAVELYRHSFRVQL
SPGIGTRLTCDIPQVKPLNIVIKLVECNKGPVAKLSDSPGKTIChDKAFVRLRKAPD
LPHIKKAS"
7775. .10387
/gene="pepN"
7775. .10387
/gene="pepN"
/note="ORF ID:0220#2
similar to SwissProt Accession Number P04825"
Query Match 100.0%; Score 1191; DB 1; Length 11470;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCCGATCTTGGGCTTGGCCGATCTGTTT 60
|||||
Db 2862 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCCGATCTTGGGCTTGGCCGATCTGTTT 2803
|||||
QY 61 CGTGCCGATGAACGTCCTCCGCGCAAAATTAACCTCGGATTGGTGTCTATATAAGATGAGACG 120
|||||
Db 2802 CGTGCCGATGAACGTCCTCCGCGCAAAATTAACCTCGGATTGGTGTCTATATAAGATGAGACG 2743
|||||
QY 121 GGCAAAACCCCGGTACTACGACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
|||||
Db 2742 GGCAAAACCCCGGTACTACGACGCGTGAAGAGGCTGAACAGTATCTGCTCGAAAAATGAA 2693
|||||
QY 181 ACCACCAAAATTAACCTCGGCGATTGACGCGATCCCTGGAATTTTGGTCGCTCACTCAGGAA 240
|||||
```

|||||
2682 ACCACAAAATTAACCTCGCATTTGACGGCATCCCTGAATTTGGTCTGCACTCAGGAA 2623
QY CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAAGTGTCTGCAGCGCACAGACT 300
Db CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAAGTGTCTGCAGCGCACAGACT 2563
QY CCGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACACAGGTT 360
Db CCGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTTGGCAAAAATACACAGGTT 2503
QY AAGCGTGTGTGGTAGCAACCCAAAGCTGCGCAACCAATGAAGAGCGTCTTAACTCTGCA 420
Db AAGCGTGTGTGGTAGCAACCCAAAGCTGCGCAACCAATGAAGAGCGTCTTAACTCTGCA 2443
QY GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAAATCACACTCTTGACTTCGAT 480
Db GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAAATCACACTCTTGACTTCGAT 2383
QY GCACTGATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTCTGTTCCATGGCTGC 540
Db GCACTGATTAAAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTCTGTTCCATGGCTGC 2323
QY TGGCATAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAACACACTGGCACAACTC 600
Db TGGCATAACCAACCGGTATCGACCCCTACGCTGGAACAATGGCAACACACTGGCACAACTC 2263
QY TCGGTTGAGAAAGCTGTTACGCTGTTGACTTTCGCTTACAGGGTTTTGGCCGCTGGT 660
Db TCGGTTGAGAAAGCTGTTTACGCTGTTTGACTTCGCTTTACAGGGTTTTGGCCGCTGGT 2203
QY CTGGAAGAAGTCTGAAGGACTCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 720
Db CTGGAAGAAGTCTGAAGGACTCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 2143
QY GCCAGTTCTTACTCTAAAACTTTGGCTCTGTAACAGAGCGTGTGGCGCTTGTACTCTG 780
Db GCCAGTTCTTACTCTAAAACTTTGGCTCTGTAACAGAGCGTGTGGCGCTTGTACTCTG 2083
QY GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCAAAATGAAGCGGCGATTCCG 840
Db GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCAAAATGAAGCGGCGATTCCG 2023
QY GCTAACTACTTAACCCACAGACACAGCGCTTCTGTTGTCACCATCTCGAGCAAC 900
Db GCTAACTACTTAACCCACAGACACAGCGCTTCTGTTGTCACCATCTCGAGCAAC 1963
QY GATGCGTTACGTGCGATTGGGAACAAGAGCTGACTGATATGGCCAGCGTATTTCAGCGT 960
Db GATGCGTTACGTGCGATTGGGAACAAGAGCTGACTGATATGGCCAGCGTATTTCAGCGT 1903
QY ATGCGTCAGTTGTTGTCGTAACGCTGCAAGAAAAGCGCAACCCGCGACTTCAGCTTT 1020
Db ATGCGTCAGTTGTTGTCGTAACGCTGCAAGAAAAGCGCAACCCGCGACTTCAGCTTT 1843
QY ATCATCAACAGAACGCGATGTTCTCTCAGTGGCCCTGACAAAGAACAAAGTGTCTCGT 1080
Db ATCATCAACAGAACGCGATGTTCTCTCAGTGGCCCTGACAAAGAACAAAGTGTCTCGT 1783
QY CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCTGAAATGTCGCGGATG 1140
Db CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCTGAAATGTCGCGGATG 1723
QY ACACGAGATAACATGGCTCCGCTGTGCAAGCGATTTGTGCGAGTGTCTGTA 1191
Db ACACGAGATAACATGGCTCCGCTGTGCAAGCGATTTGTGCGAGTGTCTGTA 1672

RESULT 9
AX370270/c
LOCUS 14759 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 87 from Patent WO0170776.
ACCESSION AX370270

AX370270.1 GI:18857433
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1
Levy, S.B., Barbosa, T.M. and Alekshun, M.N.
Nimr compositions and their methods of use
Patent: WO 0170776-A 87 27-SEP-2001;
TRUSTEES OF TUFTS COLLEGE (US)
FEATURES
source
1. 14759
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 1191; DB 6; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGAAGAACTTACCGCGCTCTGCGACCCCGATTTCTGGCCTGGCGGATCTGTTT 60
Db 2755 ATGTTTGAAGAACTTACCGCGCTCTGCGACCCCGATTTCTGGCCTGGCGGATCTGTTT 2696
QY 61 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGACG 120
Db 2695 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGACG 2636
QY 121 GGCAAAACCCCGTACTGACGACGCTGAAAAGGCTGAAACAGTATCTGCTCGAAAATGAA 180
Db 2635 GGCAAAACCCCGTACTGACGACGCTGAAAAGGCTGAAACAGTATCTGCTCGAAAATGAA 2576
QY 181 ACCACAAAATTAACCTCGGCAATGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGAA 240
Db 2575 ACCACAAAATTAACCTCGGCAATGACGCGATCCCTGAAATTTGGTCTGCTGCACTCAGAA 2516
QY 241 CTGCTGTTGTAAGTAGCGCCCTGATCAATGACAAAAGTGTCTCCACGGCACAGACT 300
Db 2515 CTGCTGTTGTAAGTAGCGCCCTGATCAATGACAAAAGTGTCTCCACGGCACAGACT 2456
QY 301 CCGGGGGCACTGGCGCACTTACGCTGGCTTTCGCGATTTCTTGGCAAAAATACACAGGTT 360
Db 2455 CCGGGGGCACTGGCGCACTTACGCTGGCTTTCGCGATTTCTTGGCAAAAATACACAGGTT 2396
QY 361 AAGCGTGTGTGGTAGCAACCCAAAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 420
Db 2395 AAGCGTGTGTGGTAGCAACCCAAAGCTGCGCAACCAATGAAGAGCGTCTTTAACTCTGCA 2336
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAAATCACACTCTTGACTTCGAT 480
Db 2335 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAAATCACACTCTTGACTTCGAT 2276
QY 481 GCACCTGATTAAAGCCTGGAATGAAGCTCAGGCTGGCGACCTAGTGTCTTCCATGGCTGC 540
Db 2275 GCACCTGATTAAAGCCTGGAATGAAGCTCAGGCTGGCGACCTAGTGTCTTCCATGGCTGC 2216
QY 541 TGCATAAACCCAAACCGGTATCGACCCCTACGCTGGAACCAATGGCAACACATGGCAACAATC 600
Db 2215 TGCATAAACCCAAACCGGTATCGACCCCTACGCTGGAACCAATGGCAACACATGGCAACAATC 2156
QY 601 TCGGTTGAGAAAGCTGTTTACCGCTGTTTGAAGTTCGCTTACAGGGTTTTGGCCGCTGGT 660
Db 2155 TCGGTTGAGAAAGCTGTTTACCGCTGTTTGAAGTTCGCTTACAGGGTTTTGGCCGCTGGT 2096
QY 661 CTGGAAGAAGTCTGAAGGACTCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 720
Db 2095 CTGGAAGAAGTCTGAAGGACTCGCGCTTTTCGCGCTATGATGAAGAGCTGATTGTT 2036
QY 721 GCCAGTTCTTACTCTAAAACTTTGGCCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG 780
Db 2035 GCCAGTTCTTACTCTAAAACTTTGGCCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG 1976

QY 781 GTTGTGCGCAGCTGAACCGTTGATCGGCAATTCAGCCAAATGAAGCGCGATTCCG 840
DB 1975 GTTGTGCGCAGCTGAACCGTTGATCGGCAATTCAGCCAAATGAAGCGCGATTCCG 1916
QY 841 GCTAACTACTCTAACCCACAGCAGCGCGCTTCTGTTGTGCGCACCACTCTGAGCAAC 900
DB 1915 GCTAACTACTCTAACCCACAGCAGCGCGCTTCTGTTGTGCGCACCACTCTGAGCAAC 1856
QY 901 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
DB 1855 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 1796
QY 961 ATGCGTCAGTTGTTCTGTAATAGCTGCGAGAAAAGGCGCAACCGCGACTTCAGCTTT 1020
DB 1795 ATGCGTCAGTTGTTCTGTAATAGCTGCGAGAAAAGGCGCAACCGCGACTTCAGCTTT 1736
QY 1021 ATCATCAAAACAGACGGCATGTTCTCTTCAGTGGCTGACAAAAGCAACAGTGTGCGT 1080
DB 1735 ATCATCAAAACAGACGGCATGTTCTCTTCAGTGGCTGACAAAAGCAACAGTGTGCGT 1676
QY 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCGTAAATGTGCGCGGATG 1140
DB 1675 CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCGTAAATGTGCGCGGATG 1616
QY 1141 ACACGAGTAACATGCTCGCTGCGGAAGCAATTTGTCAGTGTCTGTAA 1191
DB 1615 ACACGAGTAACATGCTCGCTGCGGAAGCAATTTGTCAGTGTCTGTAA 1565

RESULT 10
D90730/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

D90730 20604 bp DNA linear BCT 25-DEC-2002
Escherichia coli K12 genomic DNA. (20.9 - 21.3 min).
D90730 AB001340
D90730.1 GI:1651442
Complete and shotgun sequencing; msaB; ycaH; kdsB; ycbC; smtA;
mukF; kicB; mukE; kicA; mukB; ycbB; aspC; ompF; tolF; cmlB; coa;
cry; asnS; tss.
Escherichia coli K12
Escherichia coli K12
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

1 Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,
Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K.,
Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T.,
Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishio, H.,
Nishio, Y., Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K.,
Wada, C., Yamamoto, Y., Yano, M. and Horiuchi, T.
A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map
DNA Res. 3 (3), 137-155 (1996)
8905232

2 Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T.,
Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K.,
Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M.,
Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H.,
Motomura, K., Nakamura, Y., Nishio, H., Nishio, Y., Oshima, T.,
Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
Yamamoto, Y. and Yano, M.
The systematic sequencing of the Escherichia coli genome in Japan
Unpublished
3 (bases 1 to 20604)
Mori, H.

Direct Submission
Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science
and Technology, Res. & Edu. Center for Genetic Info.; 8916-5
Takayama, Ikoma, Nara 630-01, Japan
[E-mail:hmori@ctc.aist-nara.ac.jp, Tel:81-7437-2-5660,
Fax:81-7437-2-5669]
Collaboration Information:

Project:
The Japan E.coli genome DNA sequencing project
Group:
The Japan E.coli genome DNA sequencing group
Members: (1995.4 - 1996.3)
Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,
Horiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S.,
Itoh, T., Kanai, K., Kasai, H., Kaishimoto, K., Kim, S.,
Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K.,
Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K.,
Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N.,
Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
Yamamoto, Y. and Yano, M.
Headed by:
Name: Takashi Horiuchi
Address: National Institute of Basic Biology, Okazaki, 444, Japan
E-mail: kishori@nibb.ac.jp
Information operator:
Name: Hirotsada Mori
Address: NARA Institute of Science and Technology,
Ikoma, 630-01, Japan
E-mail: hmori@ctc.aist-nara.ac.jp
URL:
The Japan E. coli genome database
http:bsw3.aist-nara.ac.jp.

FEATURES
source
Location/Qualifiers
1..20604
/organism="Escherichia coli K12"
/mol_type="genomic DNA"
/strain="K12"
/db_xref="taxon:83333"
/map="20.9-21.3 min"
/clone="Kohara clone #218"
/note="Nucleotide position 968194-988797 from the
initiation site of ThrA (0 min.).
This clone is from Kohara lambda miniset library."

gene
1..723
/gene="msaB"
CDS
1..723
/gene="msaB"
/note="ORF ID:O217#11
similar to PIR Accession Number S27998"
/codon_start=1
/translation="EFRNVTFYPGRDVPAALRNILKIPAGTKVALVGRSGSKSTIA
SLITRFYDIDEGEILMDGHDRLREYTLASLRNOVALVSONVHLFNDTVANNIAYARTEQ
YSREQIEAAMAYAMDFINKMDNGLDTVIGNGVLLSGQRIARALLRDSFIL
LDEATSLDTESERAIOALDELQKNRTSLVIAHRLSTIEKADEIVVVEDGVIVERG
THNDLLEHRGVYQAHLKQWFGQ"
720..1706
/gene="ycalH"
720..1706
/gene="ycalH"
/note="ORF ID:O218#1
similar to PIR Accession Number S27999"
/codon_start=1
/translation="Hypothetical protein E (msaB 3' region)"
/protein_id="BAA35661.1"
/db_xref="GI:4062487"
720..1706
/gene="ycalH"

gene
720..1706
CDS
720..1706
/gene="ycalH"
/note="ORF ID:O218#1
similar to PIR Accession Number S27999"
/codon_start=1
/translation="Hypothetical protein E (msaB 3' region)"
/protein_id="BAA35661.1"
/db_xref="GI:4062487"
720..1706
/gene="ycalH"

gene
720..1706
CDS
720..1706
/gene="ycalH"
/note="ORF ID:O218#1
similar to PIR Accession Number S27999"
/codon_start=1
/translation="Hypothetical protein E (msaB 3' region)"
/protein_id="BAA35661.1"
/db_xref="GI:4062487"
720..1706
/gene="ycalH"

gene
720..1706
CDS
720..1706
/gene="ycalH"
/note="ORF ID:O218#1
similar to PIR Accession Number S27999"
/codon_start=1
/translation="Hypothetical protein E (msaB 3' region)"
/protein_id="BAA35661.1"
/db_xref="GI:4062487"
720..1706
/gene="ycalH"

gene
720..1706
CDS
720..1706
/gene="ycalH"
/note="ORF ID:O218#1
similar to PIR Accession Number S27999"
/codon_start=1
/translation="Hypothetical protein E (msaB 3' region)"
/protein_id="BAA35661.1"
/db_xref="GI:4062487"
720..1706
/gene="ycalH"

QY	841	GCTAACTACTTAACCCACGACACGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	900
Db	17223	GCTAACTACTTAACCCACGACACGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC	17164
QY	901	GATGCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT	960
Db	17163	GATGCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT	17104
QY	961	ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTTCAGCTTT	1020
Db	17103	ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAAGGCGCAAAACCGCACTTTCAGCTTT	17044
QY	1021	ATCATCAACAGAACGGCATGTTCTCTTCAGTGGGCTGACAAAAGAACAAAGTGTGCGT	1080
Db	17043	ATCATCAACAGAACGGCATGTTCTCTTCAGTGGGCTGACAAAAGAACAAAGTGTGCGT	16984
QY	1081	CTGCGGAAGAGTTTGGCGGTATATGCGTTGCTTCTGTTGCGTAAATGTGGCCGGGATG	1140
Db	16983	CTGCGGAAGAGTTTGGCGGTATATGCGTTGCTTCTGTTGCGTAAATGTGGCCGGGATG	16924
QY	1141	ACACCATATAACATGCTCGCTGTGCGAAGCGAATTGTGGCAGTGTCTGTAA	1191
Db	16923	ACACCATATAACATGCTCGCTGTGCGAAGCGAATTGTGGCAGTGTCTGTAA	16873

RESULT 11
U00096 09/c
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000

U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675
Continuation (10 of 47) of U00096 from base 900001 (U00096 Escherichia coli K-12		
Query Match 100.0%; Score 1191; DB 1; Length 110000;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGTTTGAGAACATTACCGCGCTCTCGCGACCGGATTTCTGGCGCTGGCGGATCTGTTT 60
Db	84932	ATGTTTGAGAACATTACCGCGCTCTCGCGACCGGATTTCTGGCGCTGGCGGATCTGTTT 84873
QY	61	CGTCCGATGAACGTCGCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db	84872	CGTCCGATGAACGTCGCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 84813
QY	121	GGCAAAACCCCGGTACTGACGAGCGTGAAAAGGCTGAAACAGTATCTGCTCGAAAAATGAA 180
Db	84812	GGCAAAACCCCGGTACTGACGAGCGTGAAAAGGCTGAAACAGTATCTGCTCGAAAAATGAA 84753
QY	181	ACCACCAAAATTACTCTGGCATTCAGCGATCCCTGAAATTTGGTGCCTGCACTCAGGAA 240
Db	84752	ACCACCAAAATTACTCTGGCATTCAGCGATCCCTGAAATTTGGTGCCTGCACTCAGGAA 84693
QY	241	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGACT 300
Db	84692	CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCGTCGACGGCACAGACT 84633
QY	301	CGGCGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATATACAGCGTT 360
Db	84632	CGGCGGGGCACTGGCGCACTACGCGTGGCTGCCGATTTCTGGCAAAAATATACAGCGTT 84573
QY	361	AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCATAGAGCGCTTTTAACTCTGCA 420
Db	84572	AAGCGTGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCATAGAGCGCTTTTAACTCTGCA 84513
QY	421	GGTCTGGAAGTTGCTGAATACGCTTATTATGATGCGGAAAAATCACACTTTTGACTTCGAT 480
Db	84512	GGTCTGGAAGTTGCTGAATACGCTTATTATGATGCGGAAAAATCACACTTTTGACTTCGAT 84453
QY	481	GCACTGATTAACAGCTGAACTGAAGCTCAGGCTGGCGACAGTAGTGTCTTCCATGGCTGC 540
Db	84452	GCACTGATTAACAGCTGAACTGAAGCTCAGGCTGGCGACAGTAGTGTCTTCCATGGCTGC 84393
QY	541	TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGCAACAAATGGCAAACTGGGCACAACTC 600
Db	84392	TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGCAACAAATGGCAAACTGGGCACAACTC 84333
QY	601	TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTTGGCCCGTGGT 660
Db	84332	TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGACTTTCGCTTACAGGGTTTTTGGCCCGTGGT 84273
QY	661	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTGGCGCTATGCAATAAAGAGCTGATTTT 720
Db	84272	CTGGAAGAAGATGCTGAAGGACTCGCGCTTTTGGCGCTATGCAATAAAGAGCTGATTTT 84213
QY	721	GCCAGTTCCTACTCTAAAAAATTTGGCCCTGTACAAACGAGCGGTGTGGCGCTTGACTCTG 780
Db	84212	GCCAGTTCCTACTCTAAAAAATTTGGCCCTGTACAAACGAGCGGTGTGGCGCTTGACTCTG 84153
QY	781	GTTGCTGCCGACAGTGAACACCGTTGATTCGGCGATTCAGGCCAAATGAAAGCGGCGATTTCGC 840
Db	84152	GTTGCTGCCGACAGTGAACACCGTTGATTCGGCGATTCAGGCCAAATGAAAGCGGCGATTTCGC 84093
QY	841	GCTAACTACTTAACCCACGACACGCGCGCTTCTGTGTTTGGCAACATCTCTGAGCAAC 900
Db	84092	GCTAACTACTTAACCCACGACACGCGCGCTTCTGTGTTTGGCAACATCTCTGAGCAAC 84033
QY	901	GATGCGTTACGTGCGATTTGGGAACAAGAGCTGATGATATGCGCCAGCGTATTCAGCGT 960

Db	84032	GATGCGTTACGTCGGATTTCGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT	83973
Qy	961	ATCGCTCAGTTGTTTCGTCAATACGCTGCAGGAAAAAGCGCAAAACCGGACTTTCAGCTTT	1020
Db	83972	ATGCGTCAGTTGTTTCGTCAATACGCTGCAGGAAAAAGCGCAAAACCGGACTTTCAGCTTT	83913
Qy	1021	ATCATCAACAGAACGGCATGTTCTCCTTCAGTGGCCTGCACAAAGAAACAAGTGTGCGT	1080
Db	83912	ATCATCAACAGAAACGGCATGTTCTCCTTCAGTGGCCTGCACAAAGAAACAAGTGTGCGT	83853
Qy	1081	CTGCGCAAGAGTTTGCGGTATATGCGGTGCTTCTGCTCGGCTAAATGTGCGCGGATG	1140
Db	83952	CTGCGCGAAGAGTTTGCGGTATATGCGGTGCTTCTGCTCGGCTAAATGTGCGCGGATG	83793
Qy	1141	ACACCAAGATAACATGGCTCGCTGTGCGAAGCAATTGTGGCAAGTGTCTGTA	1191
Db	83792	ACACCAAGATAACATGGCTCGCTGTGCGAAGCAATTGTGGCAAGTGTCTGTA	83742

RESULT 12

AE005174 11/C

WP_COMMENT

Sequence

```

sequence splic into 38 fragments LOCUS AE005174 ACCESSION AE005174
      Fragment Name      Begin      End

```

Log ₁₀	1	1100001	2100001	3100001	4100001	5100001	6100001	7100001	8100001	9100001	10100001	11100001	12100001	13100001	14100001	15100001	16100001	17100001	18100001	19100001	20100001	21100001	22100001	23100001	24100001	25100001	26100001	27100001	28100001	29100001	30100001	31100001	32100001	33100001	34100001	35100001	36100001	37100001	38100001	39100001	40100001	4100001	42100001	43100001	44100001	45100001	46100001	47100001	48100001	49100001
AE005174_00	AE005174_01	AE005174_02	AE005174_03	AE005174_04	AE005174_05	AE005174_06	AE005174_07	AE005174_08	AE005174_09	AE005174_10	AE005174_11	AE005174_12	AE005174_13	AE005174_14	AE005174_15	AE005174_16	AE005174_17	AE005174_18	AE005174_19	AE005174_20	AE005174_21	AE005174_22	AE005174_23	AE005174_24	AE005174_25	AE005174_26	AE005174_27	AE005174_28	AE005174_29	AE005174_30	AE005174_31	AE005174_32	AE005174_33	AE005174_34	AE005174_35	AE005174_36	AE005174_37	AE005174_38	AE005174_39	AE005174_40	AE005174_41	AE005174_42	AE005174_43	AE005174_44	AE005174_45	AE005174_46	AE005174_47	AE005174_48		

[illegible]

QY 901 GATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTTCAGCGT 960
|||||
Db 103503 GATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCAGCGTATTTCAGCGT 103444
|||||
QY 961 ATGCGTCAGTTGTTCTCAATACGCTGCGAGGAAAAAGCGCAAAACCGCGACTTCAGCTTT 1020
|||||
Db 103443 ATGCGTCAGTTGTTCTCAATACGCTGCGAGGAAAAAGCGCAAAACCGCGACTTCAGCTTT 103384
|||||
QY 1021 ATCATCAAAACAGACCGCATGTTCTCTTCAGTGGCTGACAAAGAAACAGTGTGCGT 1080
|||||
Db 103383 ATCATCAAAACAGACCGCATGTTCTCTTCAGTGGCTGACAAAGAAACAGTGTGCGT 103324
|||||
QY 1081 CTGCGGAAGAGTTTGGCGTATATGCGTTGCTTCTGCTGCGTAAATGTGGCCGGATG 1140
|||||
Db 103323 CTGCGGAAGAGTTTGGCGTATATGCTGTCTTCTGCTGCGTAAACGTTGGCCGGATG 103264
|||||
QY 1141 ACACCAGATAACATGCTCCGCTGTGCGAAGCGATTGTGGCAGTGTCTGTAA 1191
|||||
Db 103263 ACACCAGATAACATGCTCCGCTGTGCGAAGCGATTGTGGCAGTGTCTGTAA 103213
|||||

RESULT 13

AE005174_12/c

WPCOMMENT

Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name	Begin	End
AE005174_00	1	110000
AE005174_01	100001	210000
AE005174_02	200001	310000
AE005174_03	300001	410000
AE005174_04	400001	510000
AE005174_05	500001	610000
AE005174_06	600001	710000
AE005174_07	700001	810000
AE005174_08	800001	910000
AE005174_09	900001	1010000
AE005174_10	1000001	1110000
AE005174_11	1100001	1210000
AE005174_12	1200001	1310000
AE005174_13	1300001	1410000
AE005174_14	1400001	1510000
AE005174_15	1500001	1610000
AE005174_16	1600001	1710000
AE005174_17	1700001	1810000
AE005174_18	1800001	1910000
AE005174_19	1900001	2010000
AE005174_20	2000001	2110000
AE005174_21	2100001	2210000
AE005174_22	2200001	2310000
AE005174_23	2300001	2410000
AE005174_24	2400001	2510000
AE005174_25	2500001	2610000
AE005174_26	2600001	2710000
AE005174_27	2700001	2810000
AE005174_28	2800001	2910000
AE005174_29	2900001	3010000
AE005174_30	3000001	3110000
AE005174_31	3100001	3210000
AE005174_32	3200001	3310000
AE005174_33	3300001	3410000
AE005174_34	3400001	3510000
AE005174_35	3500001	3610000
AE005174_36	3600001	3710000
AE005174_37	3700001	3810000
AE005174_38	3800001	3910000
AE005174_39	3900001	4010000
AE005174_40	4000001	4110000
AE005174_41	4100001	4210000
AE005174_42	4200001	4310000
AE005174_43	4300001	4410000
AE005174_44	4400001	4510000
AE005174_45	4500001	4610000
AE005174_46	4600001	4710000

AE005174_47 4700001 4810000
AE005174_48 4800001 4910000
AE005174_49 4900001 5010000
AE005174_50 5000001 5110000
AE005174_51 5100001 5210000
AE005174_52 5200001 5310000
AE005174_53 5300001 5410000
AE005174_54 5400001 5510000
AE005174_55 5500001 5528445
Continuation (13 of 56) of AE005174 from base 1200001 (AE005174 Escherichia coli O157:H-

Query Match 98.3%; Score 1170.2; DB 1; Length 110000;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1178; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAACATTACGCGCGCTCTGCCGACCCGATTCCTGGGCTGGCCGATCTGTTT 60
Db 4403 ATGTTTGAGAACATTACGCGCGCTCTGCCGACCCGATTCCTGGGCTGGCCGATCTGTTT 4344
QY 61 CGTGGCGATGAACGTCCTGGGCAAAATTAACTTCGGGATTTGGTGTCTATAAAGATGAGCG 120
Db 4343 CGTGGCGATGAACGTCCTGGGCAAAATTAACTTCGGGATTTGGTGTCTATAAAGATGAGCG 4284
QY 121 GGCACAAACCCGGTACTGACGAGCGTGAACAAAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
Db 4283 GGCACAAACCCGGTACTGACGAGCGTGAACAAAGGCTGAACAGTATCTGCTCGAAAAATGAA 4224
QY 181 ACCACAAAAATTACCTGGCATTTGACGGCATCCCTGAATTTGGTGGCTGCACTCAGGAA 240
Db 4223 ACCACAAAAATTACCTGGCATTTGACGGCATCCCTGAATTTGGTGGCTGCACTCAGGAA 4164
QY 241 CTGCTGTTTGGTAAAGGTAGGCCCTGATCAATGACAAACGCTGCGACGGCACAGACT 300
Db 4163 CTGCTGTTTGGTAAAGGTAGGCCCTGATCAATGACAAACGCTGCGACGGCACAGACT 4104
QY 301 CCGGGGGGCACTGGCGCACTACGCGTGGCTGCGGATTTCTCGCAAAAAAATACAGCGTT 360
Db 4103 CCGGGTGGCACTGGCGCACTACGCATAGCTGCGGATTTCTGGCAAAAAAATACAGCGTT 4044
QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 4043 AAGCGAGTGTGGGTGAGCAACCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 3984
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAATACACACTCTGACTTCGAT 480
Db 3983 GATCTGGAAGTTCGTGAATACGCTTATTATGATGCGGAAACCAACCCCTGACTTCGAT 3924
QY 481 GCACCTGATTACAGCGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGCTTCCATGCTGC 540
Db 3923 GCACCTGATTACAGCGCTGAACGAGCTCAGGCTGGCGAGCTAGTGTGCTTCCATGCTGC 3864
QY 541 TGCCATAACCCAAACCGGTATCGACCCCTAGCTGGGAACAATGGCAACACATGGGCACAATC 600
Db 3863 TGCCACAACCCAAACCGGTATCGACCCCTAGCTGGGAACAATGGGCAGACACTGGGCACAATC 3804
QY 601 TCCGTTGGAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGCTTTTGGCCCGTGGT 660
Db 3803 TCCGTTGGAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGCTTTTGGCCCGTGGT 3744
QY 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGGCGCTATGCATAAAGAGCTGATTGTT 720
Db 3743 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGGCGCTATGCATAAAGAGCTGATTGTT 3684
QY 721 GCCAGTTCCTACTCTAAATAATTTTCGGCGCTGTAACAAGCGGTGTTGGCGCTTGACTCTG 780
Db 3683 GCCAGTTCCTACTCTAAATAATTTTCGGCGCTGTAACAAGCGGTGTTGGCGCTTGACTCTG 3624
QY 781 GTTGTGCGGACAGTGAACCGTTCATCGCGATTCAGCCAAATGAACGGCGATTTCGC 840
Db 3623 GTTGTGCGGACAGTGAACCGTTCATCGCGATTCAGCCAAATGAACGGCGATTTCGC 3564
QY 841 GCTAACTACTCTAACCCACCAGCACCGCGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
|||||

Db 3563 GCTAACTACTCTAACCCACCAGCACACGGCGCTTCTGTGTTGCCACCATCTCTGAGCAAC 3504

Qy 901 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 960

Db 3503 GATGCGTTACGTGCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACGCGT 3444

Qy 961 ATGCGTCAGTTGTTGCTCAATACGCTGCGAGAAAGCGCAACCCGCGACTTCAGCTTTT 1020

Db 3443 ATGCGTCAGTTGTTGCTCAATACGCTGCGAGAAAGCGCAACCCGCGACTTCAGCTTTT 3384

Qy 1021 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTCTGCGT 1080

Db 3383 ATCATCAACAGACGCGCATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTCTGCGT 3324

Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTTGGCGGGATG 1140

Db 3323 CTGCGCAAGAGTTTGGCGTATATGCTGTTGCTTCTGCTGCGTAAACGTTGGCGGGATG 3264

Qy 1141 ACACCGATTAACATGGCTCCGCTGTGCGAAGCGATTTGTCAGTGTGTA 1191

Db 3263 ACACCGATTAACATGGCTCCGCTGTGCGAAGCGATTTGTCAGTGTGTA 3213

RESULT 14
BA000007_11/c
WPCOMMENT
Sequence split into 55 fragments LOCUS BA000007 Accession BA000007

Fragment Name	Begin	End
BA000007_00	1	110000
BA000007_01	100001	210000
BA000007_02	200001	310000
BA000007_03	300001	410000
BA000007_04	400001	510000
BA000007_05	500001	610000
BA000007_06	600001	710000
BA000007_07	700001	810000
BA000007_08	800001	910000
BA000007_09	900001	1010000
BA000007_10	1000001	1110000
BA000007_11	1100001	1210000
BA000007_12	1200001	1310000
BA000007_13	1300001	1410000
BA000007_14	1400001	1510000
BA000007_15	1500001	1610000
BA000007_16	1600001	1710000
BA000007_17	1700001	1810000
BA000007_18	1800001	1910000
BA000007_19	1900001	2010000
BA000007_20	2000001	2110000
BA000007_21	2100001	2210000
BA000007_22	2200001	2310000
BA000007_23	2300001	2410000
BA000007_24	2400001	2510000
BA000007_25	2500001	2610000
BA000007_26	2600001	2710000
BA000007_27	2700001	2810000
BA000007_28	2800001	2910000
BA000007_29	2900001	3010000
BA000007_30	3000001	3110000
BA000007_31	3100001	3210000
BA000007_32	3200001	3310000
BA000007_33	3300001	3410000
BA000007_34	3400001	3510000
BA000007_35	3500001	3610000
BA000007_36	3600001	3710000
BA000007_37	3700001	3810000
BA000007_38	3800001	3910000
BA000007_39	3900001	4010000
BA000007_40	4000001	4110000
BA000007_41	4100001	4210000
BA000007_42	4200001	4310000
BA000007_43	4300001	4410000
BA000007_44	4400001	4510000

BA000007_45	4500001	4610000
BA000007_46	4600001	4710000
BA000007_47	4700001	4810000
BA000007_48	4800001	4910000
BA000007_49	4900001	5010000
BA000007_50	5000001	5110000
BA000007_51	5100001	5210000
BA000007_52	5200001	5310000
BA000007_53	5300001	5410000
BA000007_54	5400001	5498450

Continuation (12 of 55) of BA000007 from base 1100001 (BA000007 Escherichia coli Q157:1)

Query Match 98.3%; Score 1170.2; DB 1; Length 110000;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1178; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1	ATGTTTGGAGAACATTACCGCGCTCTGCGACCCGATTCCTGGCCCTGGCGGATCTCTTT	60
Db 15180	ATGTTTGGAGAACATTACCGCGCTCTGCGACCCGATTCCTGGCCCTGGCGGATCTCTTT	15121
Qy 61	CGTGCCGATGAACGTCCTCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	120
Db 15120	CGTGCCGATGAACGTCCTCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG	15061
Qy 121	GGCAAAACCCCGTACTGACACGCTGAAGGCTGAACAGTATCTGCTCGAAATGAA	180
Db 15060	GGCAAAACCCCGTACTGACACGCTGAAGGCTGAACAGTATCTGCTCGAAATGAA	15001
Qy 181	ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAAATTTGGTGCCTGCACCTCAGAA	240
Db 15000	ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAAATTTGGTGCCTGCACCTCAGAA	14941
Qy 241	CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAGCTGCTCCAGCGGCACAGACT	300
Db 14940	CTGCTGTTTGGTAAAGTAGCGCCCTGATCAATGACAAAGCTGCTCCAGCGGCACAGACT	14881
Qy 301	CCGGGGGCGACTGGCGCACTACGCTGGCTGCCGATTTCTCGCAAAAAATACCAGCGTT	360
Db 14880	CCGGGGGCGACTGGCGCACTACGCTAGCTAGCTGCCGATTTCTCGCAAAAAATACCAGCGTT	14821
Qy 361	AAGCGTGTGGTGAGCAACCCAGCTGGCCGAAACCATAGAGCGCTTTAACTCTGCA	420
Db 14820	AAGCGTGTGGTGAGCAACCCAGCTGGCCGAAACCATAGAGCGCTTTAACTCTGCA	14761
Qy 421	GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT	480
Db 14760	GATCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT	14701
Qy 481	GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTGTTCCATGGCTGC	540
Db 14700	GCACTGATTAAACAGCCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTGTTCCATGGCTGC	14641
Qy 541	TGCCATAACCCAAACCGGTATCGAACCTACGCTGGAAACATGGCAACACCTGGCAACATC	600
Db 14640	TGCCATAACCCAAACCGGTATCGAACCTACGCTGGAAACATGGCAACACCTGGCAACATC	14581
Qy 601	TCGTTTGGAGAAAGGCTGTTTACCGCTGTTTACCTTCGCTTTACCAAGGTTTTCGCCGCTGT	660
Db 14580	TCCGTTGGAGAAAGGCTGTTTACCGCTGTTTACCTTCGCTTTACCAAGGTTTTCGCCGCTGT	14521
Qy 661	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAAGAGCTGATTGTT	720
Db 14520	CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAAGAGCTGATTGTT	14461
Qy 721	GCCAGTTCTTACTCTATAAACTTTTGGCTGTACAACAGCGGTGTTGGCGCTGTGACTCTG	780
Db 14460	GCCAGTTCTTACTCTATAAACTTTTGGCTGTACAACAGCGGTGTTGGCGCTGTGACTCTG	14401
Qy 781	GTTGCTGCCGACAGTCAAAACCGTTGATCGGCAATTCAGCCAAATGAAGCGGATTCGC	840
Db 14400	GTTGCTGCCGACAGTCAAAACCGTTGATCGGCAATTCAGCCAAATGAAGCGGATTCGC	14341
Qy 841	GCTAACTACTCTAAACCCACAGCACACGGCGCTTCTGTTGTTGCCACCATCTGAGCAAC	900

||||| 14340 GCTAACTACTCTAAACCACGACACGCGCTTCTGTTGTTGCCACCACTCTCTGAGCAAC 14281
Qy 901 GATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT 960
Db 14280 GATGCGTTACGTCGGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT 14221
Qy 961 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAGGCGCAACCGGACCTTCAGCTTT 1020
Db 14220 ATGCGTCAGTTGTTGCTCAATACGCTGCAGGAAAAGGCGCAACCGGACCTTCAGCTTT 14161
Qy 1021 ATCATCAAAACAGAACGCGCATGTTCTCTTCAGTGGCTCTGACAAAAGAACAAAGTGTGCGGT 1080
Db 14160 ATCATCAAAACAGAACGCGCATGTTCTCTTCAGTGGCTCTGACAAAAGAACAAAGTGTGCGGT 14101
Qy 1081 CTGCGGGAAGAGTTTGGCGTATATGCGGTGCTTCTGCTGCGCTGCGCTAAATGTGGCCGGGATG 1140
Db 14100 CTGCGGGAAGAGTTTGGCGTATATGCTGTTGCTGCTGCGCTAAACGTTGGCCGGGATG 14041
Qy 1141 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTGTCTGTAA 1191
Db 14040 ACACCAAGATAACATGCTCGCTGTCGGAAGCGATTGTGGCAGTGTCTGTAA 13990

RESULT 15

AE005674_09/c

WPCOMMENT

Sequence split into 46 fragments LOCUS AE005674 Accession AE005674

Fragment Name	Begin	End
AE005674_01	1	110000
AE005674_02	100001	210000
AE005674_03	200001	310000
AE005674_04	300001	410000
AE005674_05	400001	510000
AE005674_06	500001	610000
AE005674_07	600001	710000
AE005674_08	700001	810000
AE005674_09	800001	910000
AE005674_10	900001	1010000
AE005674_11	1000001	1110000
AE005674_12	1100001	1210000
AE005674_13	1200001	1310000
AE005674_14	1300001	1410000
AE005674_15	1400001	1510000
AE005674_16	1500001	1610000
AE005674_17	1600001	1710000
AE005674_18	1700001	1810000
AE005674_19	1800001	1910000
AE005674_20	1900001	2010000
AE005674_21	2000001	2110000
AE005674_22	2100001	2210000
AE005674_23	2200001	2310000
AE005674_24	2300001	2410000
AE005674_25	2400001	2510000
AE005674_26	2500001	2610000
AE005674_27	2600001	2710000
AE005674_28	2700001	2810000
AE005674_29	2800001	2910000
AE005674_30	2900001	3010000
AE005674_31	3000001	3110000
AE005674_32	3100001	3210000
AE005674_33	3200001	3310000
AE005674_34	3300001	3410000
AE005674_35	3400001	3510000
AE005674_36	3500001	3610000
AE005674_37	3600001	3710000
AE005674_38	3700001	3810000
AE005674_39	3800001	3910000
AE005674_40	3900001	4010000
AE005674_41	4000001	4110000
AE005674_42	4100001	4210000
AE005674_43	4200001	4310000
AE005674_44	4300001	4410000

AE005674_44 4400001 4510000
AE005674_45 4500001 4607203
Continuation (10 of 46) of AE005674 from base 900001 (AE005674 Shigella flexneri 2a str
Query Match 98.1%; Score 1168.6; DB 1; Length 110000;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1177; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 1 ATGTTTGAACAATTACCGCGCTCTGCCGACCCGATTCTGGCTGCGCTGCCGATCTGTTT 60
Db 70933 ATGTTTGAACAATTACCGCGCTCTGCCGACCCGATTCTGGCTGCGCTGCCGATCTGTTT 70874
Qy 61 CGTGCGGATGAACGTCCTGGCAAAAATTAACTCTGGGATTTGGTGTCTATAAAGATGAGACG 120
Db 70873 CGTGCGGATGAACGTCCTGGCAAAAATTAACTCTGGGATTTGGTGTCTATAAAGATGAGACG 70814
Qy 121 GGCAAAACCCGGTACTACACGCGTGAAAAAGGTGAAACAGTATCTGCTCGAAATGAA 180
Db 70813 GGCAAAACCCGGTACTACACGCGTGAAAAAGGTGAAACAGTATCTGCTCGAAATGAA 70754
Qy 181 ACCACCAAAAATTACCTGGCATTGACGCGATCCCTGAAATTTGGTCGCTGCATCAGGAA 240
Db 70753 ACCACCAAAAATTACCTGGCATTGACGCGATCCCTGAAATTTGGTCGCTGCATCAGGAA 70694
Qy 241 CTGCTGTTTGTAAAGGTAGCGCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 300
Db 70693 CTGCTGTTTGTAAAGGTAGCGCTGATCAATGACAAACGCTGCTCGCACGGCACAGACT 70634
Qy 301 CCGGGGGCAGCTGGCGCACTACGCTGCTGCCGATTTCTCGGCAAAAAAATACCAGCGTT 360
Db 70633 CCGGGTGGCAGCTGGCGCACTACGCTGCTGCCGATTTCTCGGCAAAAAAATACCAGCGTT 70574
Qy 361 AAGCGTGTGGGTGAGCAACCCAGCTGGCCGACCAATGAAGCGTCTTTAACTCTGCA 420
Db 70573 AAGCGTGTGGGTGAGCAACCCAGCTGGCCGACCAATGAAGCGTCTTTAACTCTGCA 70514
Qy 421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAAATCACACTCTTGACTTCGAT 480
Db 70513 GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAAATCACACTCTTGACTTCGAT 70454
Qy 481 GCACCTGATTAACAGCTGTAAGAACTCAGGCTGGCGACGTAGTGTCTCATGCTGCTG 540
Db 70453 GCACCTGATTAACAGCTGTAAGAACTCAGGCTGGCGACGTAGTGTCTCATGCTGCTG 70394
Qy 541 TGCCATTAACCAACCGGTATCGACCTAGCTGGAAACATGGAACACTGGCACACACTC 600
Db 70393 TGCCATTAACCAACCGGTATCGACCTAGCTGGAAACATGGAACACTGGCACACACTC 70334
Qy 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTTCGCTTACACAGGTTTGGCCCGTGGT 660
Db 70333 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGACTTTCGCTTACACAGGTTTGGCCCGTGGT 70274
Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGGGCTATGCATAAAGAGCTGATTGTT 720
Db 70273 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTGCGGGCTATGCATAAAGAGCTGATTGTT 70214
Qy 721 GCCAGTTCCTACTCTAAAAAATTTGGCCCTGTACAACAGAGCGTGTGGCGCTTGTACTCTG 780
Db 70213 GCCAGTTCCTACTCTAAAAAATTTGGCCCTGTACAACAGAGCGTGTGGCGCTTGTACTCTG 70154
Qy 781 GTTGTGCGGACAGTGAAACCGTTCATCGCGATTTCAGCAAAATGAAAGCGCGATTTCG 840
Db 70153 GTTGTGCGGACAGTGAAACCGTTCATCGCGATTTCAGCAAAATGAAAGCGCGATTTCG 70094
Qy 841 GCTAACTACTCTAACCCACAGCACACGCGCTTCTGTTGTTGCTCCACCAATCTCTGAGCAAC 900
Db 70093 GCTAACTACTCTAACCCACAGCACACGCGCTTCTGTTGTTGCTCCACCAATCTCTGAGCAAC 70034
Qy 901 GATGCGTTACGTGCGATTTCGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT 960
Db 70033 GATGCGTTACGTGCGATTTCGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTACAGCGT 69974
Qy 961 ATGCGTCAAGTTTCTGTCATACGCTGACAGGAAAAAGGCGCAACCGCGACTTCAGCTTT 1020

D _b	69973		ATGCGT	CAGTTGTTCGTAATACGCTGCAGGAAAAAGGTGCAAAACCGCAGATTTCAGCTTT	69914
Q _y	1021	ATCATCAAACAGNACGGCATGTTCTCTTCAGTGCCCTGACAAAAGAACAAGTCTCGGT	1080		
D _b	69913	ATCATCAAACAGAACGCCATGTTCTCTTCAGTGTCTGACAAAAGAACAAAGTCTCGGT	69854		
Q _y	1081	CTGCGGGAAGAGTTTTGGCGGTATATCGGTTGCTTCTGCTCGGTAAATGTGGCCGGGATG	1140		
D _b	69853	CTGCGCGAAGAGTTTGGCGGTATATGCTGTGTTCTTGTGTCGGTAAATGTGGCCGGGATG	69794		
Q _y	1141	ACACCAGATAACATGCTCCGCTGTGCGAAGCATTTGTGGCAGTGCTGTAA	1191		
D _b	69793	ACACCAGATAACATGCTCCGCTGTGCGAATCGATTTGGCAGTGCTGTAA	69743		

Search completed: March 15, 2006, 15:50:55
Job time : 6324 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 17:52:30 ; Search time 189 Seconds
(without alignments)

920.603 Million cell updates/sec

Title: US-10-673-786A-2

Perfect score:

Sequence: 1 MFENITAAPADPILGLADLF.....VAGMTPDNMAPLCEAIVAVL 396

Scoring table: BLOSUM62

scoring table: BDC030M02
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

100% processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 21:*

- 1: genesetp1980s:*
- 2: genesetp1990s:*
- 3: genesetp2000s:*
- 4: genesetp2001s:*
- 5: genesetp2002s:*
- 6: genesetp2003as:*
- 7: genesetp2003bs:*
- 8: genesetp2004s:*
- 9: genesetp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2045	100.0	396	1	AAP71678	Aap71678 aspC gene
2	2045	100.0	396	7	AW69553	Aw69553 Escherich
3	2045	100.0	396	2	ADB83285	AdB83285 Bacterial
4	2045	100.0	396	8	ADN18132	Adn18132 Bacterial
5	2045	100.0	396	9	ADW95404	Adw95404 Amino aci
6	1931	94.4	402	7	ABO65556	AbO65556 Klebsiell
7	1682	82.2	397	6	ABM69171	ABm69171 Photorhab
8	1654	80.9	396	8	ADS42695	AdS42695 Bacterial
9	1318	64.4	322	8	ADN17578	Adn17578 Bacterial
10	1317	64.4	423	6	ABP80003	ABp80003 N. gonorr
11	1311	64.1	397	5	AU73004	Au73004 Neisseria
12	1310	64.1	397	8	ADP08222	ADp08222 Neisseria
13	1005.5	49.2	394	8	ADS22791	AdS22791 Bacterial
14	996	48.7	396	8	ADN24844	Adn24844 Bacterial
15	996	48.7	398	8	ADN22085	Adn22085 Bacterial
16	982	48.0	397	8	ADN25771	Adn25771 Bacterial
17	982	48.0	413	7	ABO80253	ABo80253 Pseudomon
18	975	47.7	397	8	ADS26795	AdS26795 Bacterial
19	975	47.7	397	8	ADS27175	AdS27175 Bacterial
20	975	47.7	398	8	ADS26427	AdS26427 Bacterial
21	947	46.3	398	8	ADS24992	AdS24992 Bacterial
22	924	45.2	395	8	ADN26805	Adn26805 Bacterial
23	917	44.8	395	8	ADN26569	Adn26569 Bacterial
24	917	44.8	420	8	ADY07798	Ady07798 Plant fun

ALIGNMENTS

RESULT 1

RESOLUT I
AAP71678
ID AAP71678 standard; protein: 396 AA.

AA
AC
AAP71678:XX
DT 01-JAN-1980 (first entry)

agnc gene product from plasmid pME219

XX
KW alpha-amylase; feedback inhibition; amino acid synthesis;
KW composite plasmid.

xx Bacillus licheniformis.

XX
PN
W08700302-A

15--JAN-1987

XX
PF 24-JUN-1986. 86W0-IIS001353

XX
24-JUN-1985. 85JIS-00747732
DPXX
PA (NITR-) NITRASWEET CO

XX
PT Edwards MP Taylor PP Hunter MG Forburingh IG:

XX
WPT: 1987-021998/03

DR N-PSDB; AAN71109.
XX

Composite plasmids contg. multiple genes in transcriptional units -
 useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.

PS
Disclosure; Page 38; 5/99; English.
XX

This sequence may produced from a composite plasmid and used for the production of amino acids. See also AAP71053-55, AAP71107, AAP71109, CC AAP70696-97 and AAP70750. AAP70752 and AAP70754

XX S0 Sequence 396 AA:

Query Match

Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396: Conservative 0: Mismatches 0: Indels 0

Qv 1 MEENTAAPADPTILGLADLERADERPGKINLGIGVYKD

Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPVLTSVKKAQYLLENE 60
Qy 61 TTKNYLIGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLIGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Db 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFEAYQGFARGLEDAEGLRFAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFEAYQGFARGLEDAEGLRFAFAAMHKELIV 240
Qy 241 ASSYSKNFGLYNERVAGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVAGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Qy 361 LREEFGVYASGRVNVAGTMDNMAPLCEAIVAVL 396
Db 361 LREEFGVYASGRVNVAGTMDNMAPLCEAIVAVL 396

RESULT 2
AAW69553
ID AAW69553 standard; protein; 396 AA.
XX AC AAW69553;
XX DT 13-OCT-1998 (first entry)
XX DE Escherichia coli aspC protein.
XX KW Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
XX KW aspartokinase; feedback inhibition; dihydropicolinate reductase;
XX KW diaminopimelate decarboxylase; aspartate aminotransferase.
XX OS Escherichia coli.
XX PN EP854189-A2.
XX PD 22-JUL-1998.
XX PF 05-DEC-1997; 97EP-00121443.
XX PR 05-DEC-1996; 96JP-00325659.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;
XX DR WPI; 1998-379060/33.
XX DR N-PSDB; AAV40259.
XX PT Recombinant DNA autonomously replicable in coryneform bacteria - used to
PT produce L-lysine, codes for e.g. aspartokinase, dihydropicolinate
PT reductase and synthase and di.amino-pimelate decarboxylase.
XX PS Claim 6; Page 38-39; 59pp; English.
XX CC The present invention describes a recombinant DNA autonomously replicable
CC in cells of coryneform bacteria (CB), comprising a DNA sequence coding
CC for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
CC threonine is desensitised, a DNA sequence coding for a
CC dihydropicolinate reductase (DHPR), a DNA sequence coding for
CC dihydropicolinate synthase (DHPS), a DNA sequence coding for
CC diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
CC aspartate aminotransferase (AAT). The present sequence represents aspC
CC from Escherichia coli. The DNA and related products from the present

CC invention, can be used for improving L-lysine productivity by CB. The L-
CC lysine produced can be used as a fodder additive
SQ Sequence 396 AA;
Query Match 100.0%; Score 2045; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPVLTSVKKAQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLIGVYKDETKTPVLTSVKKAQYLLENE 60
Qy 61 TTKNYLIGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLIGIDGIPFGRCCTOELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120
Qy 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Db 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
Qy 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFEAYQGFARGLEDAEGLRFAFAAMHKELIV 240
Db 181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPDFEAYQGFARGLEDAEGLRFAFAAMHKELIV 240
Qy 241 ASSYSKNFGLYNERVAGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVAGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
Qy 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Db 301 DALRAIWEQELTDMRQRIORMRQLFVNTLQEKGNRDFSFIIKQNGMFSFGLTKQVLR 360
Qy 361 LREEFGVYASGRVNVAGTMDNMAPLCEAIVAVL 396
Db 361 LREEFGVYASGRVNVAGTMDNMAPLCEAIVAVL 396
RESULT 3
ADB83285
ID ADB83285 standard; protein; 396 AA.
XX AC ADB83285;
XX DT 04-DEC-2003 (first entry)
XX DE Escherichia coli aspartate aminotransferase.
XX KW aspartate aminotransferase; threonine; fermentation; enzyme.
XX OS Escherichia coli.
XX PN WO2003072786-A1.
XX PD 04-SEP-2003.
XX PF 25-FEB-2003; 2003WO-JP002067.
XX PR 27-FEB-2002; 2002RU-00104983.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;
XX DR WPI; 2003-721777/68.
XX DR N-PSDB; ADB83284.
XX PT Industrial production of L-threonine by fermentation using Escherichia
PT modified to enhance aspartate aminotransferase activity, with improved
PT productivity.
XX PS Claim 7; Page 20-21; 26pp; Japanese.

CC The invention relates to a bacterium belonging to the genus *Escherichia*
CC which is modified to enhance aspartate aminotransferase activity for the
CC production of L-threonine. The bacterium is used for the industrial
CC production of L-threonine by fermentation. This sequence corresponds to
CC the *E. coli* aspartate aminotransferase protein.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 2045; DB 7; Length 396;
Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPTLVTSVKAEQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPTLVTSVKAEQYLLENE 60

Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGALRVAADFLAKNTSV 120

Qy 121 KRVMVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Db 121 KRVMVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180

Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240

Qy 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFQSMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFQSMKAAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQIRQMRLFVNTLQEKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQIRQMRLFVNTLQEKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360

Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 4
ADN18132
ID ADN18132 standard; protein; 396 AA.

AC ADN18132;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #785.

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.
OS
XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.

(GOLD/) GOLDMAN B S.
PA
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 785; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 396 AA;

Query Match 100.0%; Score 2045; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 2e-197;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPTLVTSVKAEQYLLENE 60
Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKTPTLVTSVKAEQYLLENE 60

Qy 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGALRVAADFLAKNTSV 120
Db 61 TTKNYLGIDGIPFGRCTQELLFGKGSALINDKARTATQTPGGTGALRVAADFLAKNTSV 120

Qy 121 KRVMVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
Db 121 KRVMVSNPSPNHNKSVFNSAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180

Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240
Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGMWLPDFAYQGFARGLEDAEGLRFAAMHKLIV 240

Qy 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFQSMKAAIRANYSNPPAHGASVVATILSN 300
Db 241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFQSMKAAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQIRQMRLFVNTLQEKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360
Db 301 DALRAIWEQELTDMRQIRQMRLFVNTLQEKGNRDFSFIIKQNGMFSFSGLTKEQVLR 360

Qy 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396
Db 361 LREEFGVYAVASGRVNVAGMTPDNMAPLCEAIVAVL 396

RESULT 5

ADW95404	ADW95404 standard; protein; 396 AA.	
ID	ADW95404	
XX	ADW95404	
AC	ADW95404	
XX	ADW95404	
DT	07-APR-2005 (first entry)	
XX	07-APR-2005 (first entry)	
XX	Amino acid sequence of an Escherichia coli polypeptide.	
DE	transgenic plant; glutamic acid dehydrogenase; GDH; ECASPC;	
XX	2-oxo glutaric acid; agriculture.	
KW	Escherichia coli.	
XX	Escherichia coli.	
OS	WO2005006847-A1.	
XX	27-JAN-2005.	
PN	15-JUL-2004; 2004WO-JP010451.	
PD	17-JUL-2003; 2003JP-00198559.	
XX	(AJIN) AJINOMOTO CO INC.	
XX	Kisaka H, Miwa T, Akiyama A;	
PI	WPI; 2005-132242/14.	
DR	N-PSDB; ADW95403.	
XX	Producing plant having improved growth and yield under cultivation	
PT	conditions with decrease in nitrogen, involves expressing transduced	
PT	glutamic acid dehydrogenase gene in plant so as to increase their 2-oxo	
PT	glutamic acid contents.	
XX	Disclosure; SEQ ID NO 20; 94pp; Japanese.	
PS	The specification describes a method of producing a plant having improved	
XX	growth and yield under cultivation conditions with a decrease in	
CC	nitrogen. The method involves transducing glutamic acid dehydrogenase	
CC	(GDH) or ECASPC genes into a plant and expressing the gene so as to	
CC	increase 2-oxo glutaric acid content of the plant, or applying proline to	
CC	the foliage of the plant so as to increase the 2-oxo glutaric acid	
CC	content of the plant. The method is useful for producing a plant having	
CC	improved growth and yield under cultivation conditions with a decrease in	
CC	nitrogen. The present sequence represents a polypeptide, used in the	
CC	course of the invention.	
XX	Sequence 396 AA;	
SQ	Query Match 100.0%; Score 2045; DB 9; Length 396;	
	Best Local Similarity 100.0%; Pred. No. 2e-197;	
	Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTSVKKAQYLLENE 60	
DB	1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTSVKKAQYLLENE 60	
QY	61 TTKNYLIGDIGIPEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120	
DB	61 TTKNYLIGDIGIPEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120	
QY	121 KRVMVSNPWNHKSVMFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180	
DB	121 KRVMVSNPWNHKSVMFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180	
QY	181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGRLAFAMHKKELIV 240	
DB	181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGRLAFAMHKKELIV 240	
QY	241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSSQMAAIRANYSNPPAHGASVVATILSN 300	
DB	241 ASSYSKNFGLYNERVAGCTLVAAADSETVDRAFSSQMAAIRANYSNPPAHGASVVATILSN 300	
QY	301 DALRAIWEQELTDMRQRIQMRQLFVNTLQEKGNDRFSFIKQNGMFSFSGLTKEQVLR 360	
DB	301 DALRAIWEQELTDMRQRIQMRQLFVNTLQEKGNDRFSFIKQNGMFSFSGLTKEQVLR 360	
QY	361 LRBEFGYIVASGRVNVAGTMDNMAPLCEAIVAVL 396	
DB	361 LRBEFGYIVASGRVNVAGTMDNMAPLCEAIVAVL 396	
RESULT 6		
ABO65556		
ID	ABO65556 standard; protein; 402 AA.	
XX	ABO65556;	
XX	29-JUL-2004 (first entry)	
DT	Klebsiella pneumoniae polypeptide seqid 12073.	
DE	Recombinant expression vector; transcription regulatory element;	
XX	Klebsiella pneumoniae protein; antibacterial; Vaccine.	
KW	Klebsiella pneumoniae.	
OS	US6610836-B1.	
XX	26-AUG-2003.	
PN	27-JAN-2000; 2000US-00489039.	
PD	29-JAN-1999; 99US-0117747P.	
XX	(GENO-) GENOME THERAPEUTICS CORP.	
XX	Breton GL, Osborne M;	
PI	WPI; 2003-895346/82.	
DR	N-PSDB; ACH99107.	
XX	New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for	
PT	preparing a vaccine composition against Klebsiella pneumoniae.	
PT	Disclosure; SEQ ID NO 12073; 932pp; English.	
PS	The invention describes a new isolated nucleic acid encoding a Klebsiella	
XX	pneumoniae polypeptide. Also described are: a recombinant expression	
CC	vector comprising the nucleic acid, operably linked to a transcription	
CC	regulatory element; and a cell comprising the recombinant expression	
CC	vector. The nucleic acid is useful for preparing a vaccine composition	
CC	against Klebsiella pneumoniae. This is the amino acid sequence of a	
CC	Klebsiella pneumoniae polypeptide of the invention	
XX	Sequence 402 AA;	
SQ	Query Match 94.4%; Score 1931; DB 7; Length 402;	
	Best Local Similarity 93.4%; Pred. No. 6.9e-186;	
	Matches 370; Conservative 12; Mismatches 14; Indels 0; Gaps 0;	
QY	1 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTSVKKAQYLLENE 60	
DB	7 MFENITAAPADPILGLADLFRADRPCKINLIGVYKDETKPTVLTSVKKAQYLLENE 66	
QY	61 TTKNYLIGDIGIPEFGRCCTQELLFGKGSALINDKRAARTAQTPGGTGALRVAADFLAKNTSV 120	
DB	67 TTKNYLIGDIGIPEFGRCCTQELLFGKGNATADKRAARTAQTPGGTGALRVAADFLAKNTDV 126	
QY	121 KRVMVSNPWNHKSVMFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 180	
DB	127 KRVMVSNPWNHKSVMFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLFHGC 186	
QY	181 CHNPTGIDPTLEQWQTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGRLAFAMHKKELIV 240	
DB	187 CHNPTGIDPTLEQWQTLAQLSVEKGWLPFLDFAYOGFARGLEEDAEGRLAFAMHKKELIV 246	

QY 241 ASSYSKNFGLYNERVAGTVAADSETVDRAFQKAAIRANYSNPPAHGASVATILSN 300
 DB 247 ASSYSKNFGLYNERVAGTVAADQETVDRAFQKSVIRANYSNPPAHGASVATILSN 306
 QY 301 DALRAIWEQELTDMRQRIORMQLFVNTLOEKGANDRDFSIKQNGMFSFSLGTEQVLR 360
 DB 307 DALRAIWEQELTDMRQRIORMQLFVNTLOEKGASRDFSIQNGMFSFSLGTEQVLR 366
 QY 361 LREEFGVAVASGRVNVAGTMDNNAPLCEAIVAVL 396
 DB 367 LREEFAIVAVASGRINVAGTMDNNAPLCEAIVAVL 402

RESULT 7

ID ABM69171 standard; protein; 397 AA.
 XX
 AC ABM69171;

20-NOV-2003 (first entry)

Photorhabdus luminescens protein sequence #2268.

Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 detection; food; gene expression; plant; animal; microorganism; toxin;
 antibiotic; biopesticide; virulence factor; disease model; plague;
 whooping cough.

Photorhabdus luminescens.

WO200294867-A2.

28-NOV-2002.

07-FEB-2002; 2002WO-IB003040.

07-FEB-2001; 2001FR-00001659.

(INST) INST PASTEUR.

(CNRS) CNRS CENT NAT RECH SCI.

Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 Buchrieser C;

WPI; 2003-148459/14.

Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 useful e.g. as therapeutic antimicrobials and agricultural pesticides.

Claim 2; SEQ ID NO 2268; 1205pp; French.

The invention relates to the isolation of genes and their encoded
 proteins from Photorhabdus luminescens. The isolated sequences are
 sources of probes and primers for detecting the genome of P. luminescens
 and related species; to study polymorphisms; for gene analysis and for
 detection/amplification of the genes. Antibodies (Ab) raised against the
 polypeptides encoded by the genes are used for detection/identification
 of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 carry a gene-containing vector are used to select compounds that
 modulate, regulate, induce or inhibit expression of the genes in plants,
 animals or microorganisms other than P. luminescens and are able to alter
 response or sensitivity to toxins and antibiotics produced by P.
 luminescens. Cells transformed to express the genes are useful for
 recombinant production of the proteins, particularly toxins and
 antibacterials useful as insecticides, bactericides and fungicides. The
 genes, proteins, vectors containing the genes and Ab are also useful
 therapeutically (to treat microbial infection by bacteria or fungi that
 are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 biopesticides. Other uses of the genes and the proteins are as virulence
 factors and for identifying targets of human diseases for which P.
 luminescens is a model (particularly plague and whooping cough). This
 sequence represents one of the isolated P. luminescens proteins

XX
 SQ

Sequence 397 AA;

Query Match 82.2%; Score 1682; DB 6; Length 397;

Best Local Similarity 79.8%; Pred. No. 1e-160;
 Matches 316; Conservative 38; Mismatches 42; Indels 0; Gaps 0;

QY 1 MFENITAAPADPILGLADLFRADERPGKINLIGIVYKDBTGKTPVLTSVKKAEOYLLEN 60

DB 1 MFEKITAAPADPILGLADSFSDPTNKINLIGIVYKDBTGKTPVLTSVKKAEOYLLEN 60

QY 61 TTKNYLGIDGPEFCRTOELLFGKGSALINDKRARTATQPGTGALRVAADFLAKNTSV 120

DB 61 TTKNYLPISGLAEFGKRVTOELLFGKHDPVWTDKARTAQSPGTTGALRIAADFLAKNTNA 120

QY 121 KRVVSNFSPNHNKSVFNSAGLEVPEYAYDAENHTLDFDALINSINEAQAGDVVLFHGC 180

DB 121 KRVWISNPTWPNHNKVFNSAAGLEVCEYKYDAEKHALNFMEDLASLSEAQAQGVVLFHGC 180

QY 181 CHNPTGIDPTLEQMOTLAQLSVKGLWLPFAYOGFARGLEDEAGLRAFAAMHKLIV 240

DB 181 CHNPTGIDPTPAQWAKLAEMSAEKGLWLPFAYOGFARGLEDEAGLRAFAAMHKLIV 240

QY 241 ASSYSKNFGLYNERVAGTVAADSETVDRAFQKAAIRANYSNPPAHGASVATILSN 300

DB 241 ASSYSKNFGLYNERVAGTIVASDSDTAEKAFSQAKAIIRANYSNPPAHGASIVTILSN 300

QY 301 DALRAIWEQELTDMRQRIORMQLFVNTLOEKGANDRDFSIKQNGMFSFSLGTEQVLR 360

DB 301 EDLKAWEQELTTRERIQRMQLFVNTLOEKGAKQDFSIISQNGMFSFSLGTEQVER 360

QY 361 LREEFGVAVASGRVNVAGTMDNNAPLCEAIVAVL 396

DB 361 LRDEFGIYAVSSGRINVAGLTENNAPLCEAIVAVL 396

RESULT 8

ADS42695

ID ADS42695 standard; protein; 396 AA.

XX ADS42695;

02-DEC-2004 (first entry)

Bacterial polypeptide #21125.

Recombinant DNA construct; transformed plant; improved plant property;
 cold tolerance; heat tolerance; drought tolerance; herbicide; oomosis;
 pathogen tolerance; pest tolerance; plant disease resistance;
 cell cycle pathway modification; plant growth regulator;
 homologous recombination; seed oil yield; protein yield; carbohydrate;
 nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

DR

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
PS Claim 1; SEQ ID NO 21125; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 396 AA;

Query Match 80.9%; Score 1654; DB 8; Length 396;
Best Local Similarity 78.3%; Pred. No. 7.1e-158;
Matches 310; Conservative 34; Mismatches 52; Indels 0; Gaps 0;
QY 1 MFENITAAPADPILGLADLFRADPERGKINLGIGYKDETGTPTLTSVKKAEQYLLNE 60
DB 1 MFEKITAAPADPILGLADSPKADPRENKINLGIGYKDETGTPTLTSVKKAEFLLENE 60
QY 61 TTKNYLGDIGTPEFCRGTQELLFGKGSALINDKARTATQGTGTCALRVAADFLAKNTSV 120
DB 61 TTKNYLAISGLPEFGRVTVQELLFGNTSTIITDKGARTVQSPGTCALRTAADFIAKQTN 120
QY 121 KRVVSNPSPNHNKSVNSAGLEVREYAYDAENHTLDFDALINSLNQAQGVVLPFGC 180
DB 121 KRVWISNPTWPNHKGVSAGLEIREYQYNAEKHALDFDGLMSLSAQAGDVVLLHGC 180
QY 181 CHNPTGIDPTLEQWQTLAQLSVKGLPLFPDFAYQGFARGLEEDAGLRFAAMHKEIV 240
DB 181 CHNPTGIDPTAEQWQKLADLSAANGWLPVDFAYQGFARSLEDEDAEGLRIFTKHNEILV 240
QY 241 ASSYSKNFGLNRYRGACTLVADSETVDRAFSPQKAAIRANYSNPPAHGASVATILSN 300
DB 241 ASSYSKNFGLNRYRGACTIVATSDTAKAFSAQKSIVRTNYSNPPAHGASVVTILSN 300
QY 301 DALRAIWQELTDMRQRLORMLFVNTLOEKGANRDFSLIKONGMFSFSLGTEQVLR 360
DB 301 DEFKAEWTOELATMRERIRRMKQLFVNTLOEKGAKQDFSLISQNGMFSFSLGTEQVDR 360
QY 361 LREBFQVAVASGRVNVAGMTDDNNAPLCEAIVAVL 396
DB 361 LREBFQVAVSSGRINRAGLTLENVPLCEAIVAVL 396

RESULT 9
ADN17578
ID ADN17578 standard; protein; 322 AA.
XX
AC ADN17578;
XX

DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #231.
DE
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
XX 18-DEC-2003.
PD
XX 20-FEB-2003; 2003US-00369493.
PF
XX 21-FEB-2002; 2002US-0360039P.
PR
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 231; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC having an improved property comprising transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 322 AA;

Query Match 64.4%; Score 1318; DB 8; Length 322;
Best Local Similarity 76.4%; Pred. No. 5e-124;
Matches 246; Conservative 32; Mismatches 44; Indels 0; Gaps 0;
QY 75 GRCTQELLFGKGSALINDKARTATQGTGTCALRVAADFLAKNTSVKRVWISNPSWPNHK 134
DB 1 GRVIOELLFGSTSAIVTEKARTVQSPGTCALRTAADFIAKQNAKRVWISNPTWPNHK 60

QY 135 SVFNSAGLEVREYAYDAENHTLDFDALINSNEAQAQDVVLFHGCHNPTGIDPTLEQM 194
 DB 61 GVFSAGLEIREYNYDAEKHALNFEGLASLASEQAQDVVLLHGCHNPTGIDPTPEQM 120
 QY 195 OTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRFAAMHKELEIVASSYKNGFLYNER 254
 DB 121 OKLADLSAANGWLPVDFPAYQGFARGLEEDAEGLRFAAMHKELEIVASSYKNGFLYNER 180
 QY 255 VGACTLVAADSETVDRAFSQMAAIIIRANYSNPPAHGASVVATILSNDALRAIWEQELTDM 314
 DB 181 VGACTLVAADSDTAERAFQAKFIVRTNYSNPPAHGASVVATILSNEELKAEWIOELATM 240
 QY 315 RRIQRMQLFVNTLQEKAGNDFSFIIKONGMFSFGLTKBOVLRLREFFGVYAVASGR 374
 DB 241 RERIQRMQLFVNTLQEKAGQDFSFICAQNGMFSFGLTKBOVLRLREFFGVYAVASGR 300
 QY 375 VNVAGTTPDNMAPLCEAIVAVL 396
 DB 301 INVAGLTLENMVLCEAIVAVL 322

RESULT 10
 ABP80003
 ID ABP80003 standard; protein; 423 AA.
 XX
 AC ABP80003;
 DT 07-MAR-2003 (first entry)
 DE N. gonorrhoeae amino acid sequence SEQ ID 6536.
 KW Antibacterial; infection; vaccine; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-1B002069.
 XX
 PR 12-FEB-2001; 2001GB-00003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizza M, Massignani V, Monaci E;
 XX
 DR WPI; 2003-058415/05.
 DR N-PSDB; ABZ40973.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 PS Disclosure; Page 663; 815pp; English.
 XX

RESULT 11
 AAU73004
 ID AAU73004 standard; protein; 397 AA.
 XX
 AC AAU73004;
 DT 12-MAR-2002 (first entry)
 DE Neisseria meningitidis virulence protein #94.
 KW Meningitis; virulence; gene; antibacterial; vaccine; veterinary;
 KW infection; Gram-negative bacteria; antimicrobial.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200185772-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 08-MAY-2001; 2001WO-GB002003.
 XX
 PR 08-MAY-2000; 2000GB-00011108.
 XX
 PA (MICR-) MICROSCIENCE LTD.
 XX
 PI Tang C;
 XX
 DR WPI; 2002-066593/09.
 DR N-PSDB; AAS97289.
 XX
 PT New peptide encoded by operon including virulence genes of Neisseria
 PT meningitidis, useful as vaccine component for treating or preventing
 PT meningitis and for identifying antimicrobial drug.
 XX
 PS Claim 4; Page 381-383; 423pp; English.
 XX
 CC The invention relates to a peptide (I) encoded by an operon (II) of
 CC Neisseria meningitidis including virulence genes, or a related molecule
 CC having a 40% sequence similarity at the peptide or nucleotide level in a
 CC Gram-negative bacterium, or its functional fragment, for therapeutic or
 CC diagnostic use. (I) and (II) are useful in the manufacture of a
 CC medicament for treating or preventing a condition (e.g., meningitis)
 CC associated with infection by Neisseria or Gram-negative bacteria. The
 CC product is useful for veterinary treatment and in a screening assay for
 CC the identification of an antimicrobial drug. The vaccines have
 CC prophylactic applications. AAU72911-AAU73014 represent N. meningitidis
 CC virulence proteins of the invention

QY 62 TKNYLIDGIPFGRCTOELLFSGSALINDKRARTAOPTGGTALRVAADFALKNSTVK 121
 DB 90 TKNYLIDGVADYNEQTQILLFGKDEHIIASRAKTAQSLGGTALRIABEAKRQQLNAQ 149
 QY 122 RVVNSPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSNEAQAQDVVLFHGCH 181
 DB 150 TTIWISNPTWPNHNAIAKAVGIQDKPYRYDAAKHCLDMDGMIEDLNQAQKGDIVLLHGCC 209
 QY 182 HNPTGIDPTLEOMOTLAQLSVKGLPLDFPAYQGFARGLEEDAEGLRFAAMHKELEIVA 241
 DB 210 HNPTGIDPTPEQWETLAKLSAEKGLPLDFPAYQGFNGLEEDAYGLRVFLKHNTLELLA 269
 QY 242 SSKSNFGLYNERVGAFTLVAADSETVDRAFSQMAAIIIRANYSNPPAHGASVVATILSND 301
 DB 270 SSKSNFGLYNERVGAFTLVAADSETVDRAFSQMAAIIIRANYSNPPAHGASVVATILSND 329
 QY 302 ALRAIWEQELTDMRQIORMQLFVNTLQEKAGNDFSFIIKONGMFSFGLTKBOVLRL 361
 DB 330 DLKAQWIAELEDGRKAMQKQFVLLKAKGSTQDFDIIQNGMFSFGLTKBOVLRL 389
 QY 362 REEFGVYAVASGRVNVAGTTPDNMAPLCEAIVAV 395
 DB 390 KNEFAIYAVRSGRINVAGITDDNIDYLCESIVKV 423

Query Match 64.4%; Score 1317; DB 6; Length 423;
 Best Local Similarity 61.9%; Pred. No. 9.8e-124;
 Matches 244; Conservative 62; Mismatches 88; Indels 0; Gaps 0;
 QY 2 FENITAAPADPILGLADIUFRADEPGKINLGIYGVYKBTGKTPVLTSVKKAPQYLLENET 61
 DB 30 FKHEAAPADPILGLGEAFKAEATREPKVNLGIGVYKDSAGATPIVKAKEAKRLLESET 89

```

XX Sequence 397 AA;
SQ
Query Match      64.1%; Score 1311; DB 5; Length 397;
Best Local Similarity 61.8%; Pred. No. 3.6e-123;
Matches 244; Conservative 62; Mismatches 89; Indels 0; Gaps 0;

QY 2 FENITAPADPILGLADLFRADERPGKINLIGVYKDTGKTPVLTTSVKAEQVLLNET 61
DB 3 FKHIAPADPILGLGEAFKAETRPKVNIGVYKDSAGATPIVKAKEAKLLSESET 62

QY 62 TKNYLIGIDGIPFGRCCTQELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSVK 121
DB 63 TKNYLTIIDGVADYNAQTQILLFGKDHEIIASRRAKTAQSLGCTGALRIAFAFAKQLNAQ 122

QY 122 RVVSNPSPWPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQGDVVLFGCC 181
DB 123 TIWISNPTWPNHNAIAKAVGIQDPYRYDAAGHGLDWDGMIEDLSQAQKGDIVLLHGCC 182

QY 182 HNPTGIDTPTLEQWOTLAQLSVEKGMPLDFPAYQGFARGLEDAEGLRAPAMHKLIVA 241
DB 183 HNPTGIDTPTLEQWETLAKLSAEKGMPLDFPAYQGFNGLEBEDAYGLRVFLKHNTPELLIA 242

QY 242 SSKNFGLYNERVAGCTLVAAADSETVDRAFSPQKAAIRANYSNPPAHGASVATILSND 301
DB 243 SSKNFGLYNERVAGFTLVAADEATAARAHSQVKTIIRTLYSNPASHGANTIALVLKND 302

QY 302 ALRAIWEQELTDMRQRIQRMQLFVNTLQEKGANRDFSFIKONGMFSFGLTKEQVLR 361
DB 303 DLKAQWIAELDEMGRKAMQKQFVELLKAKGATQDFDIIIEQNGMFSFGLTPEQVDR 362

QY 362 REEFGVYVAGSRVNVAGTDPDNPAPLCEAIVAVL 396
DB 363 KNEFAIYVRSGRINVAGITDDNIDYLCESIVKVL 397

RESULT 12
ADP08222
ID ADP08222 standard; protein; 397 AA.
XX
AC ADP08222;
XX
XX ADP08222;
XX
DT 26-AUG-2004 (first entry)
XX
DE Neisseria meningitidis MC58 OMV-related membrane protein - SEQ ID 55.
KW outer-membrane vesicle; antibacterial; antiinflammatory;
KW meningococcal protein trafficking; localisation; infection; vaccine;
KW Gene therapy.
XX
OS Neisseria meningitidis MC58.
XX
PN WO2004046177-A2.
XX
PD 03-JUN-2004.
XX
PF 17-NOV-2003; 2003WO-IB006281.
XX
PR 15-NOV-2002; 2002GB-00026734.
PR 27-MAR-2003; 2003GB-00007131.
XX
XX (CHIR ) CHIRON SRL.
XX
PI Norais N, Grandi G;
XX
XX WPI; 2004-420615/39.
XX
XX New compositions having outer-membrane vesicles and proteins from
PT Neisseria meningitidis, useful in the field of meningococcal
PT biochemistry, in particular for preventing and/or treating meningococcal
PT infections.
XX
XX Claim 9; SEQ ID NO 55; 79pp; English.

```

```

XX The invention relates to a novel composition comprising outer-membrane
CC vesicles (OMV) prepared from a first strain of Neisseria meningitidis and
CC 1 or more proteins which are present in OMVs prepared from a second
CC strain of N. meningitidis, but which are not present in OMVs prepared
CC from the first strain. The composition of the invention demonstrates
CC antibacterial and antiinflammatory activities and may be useful in the
CC field of meningococcal biochemistry, in particular the trafficking and
CC localisation of meningococcal proteins, as well as in the prevention or
CC treatment of meningococcal infections, possibly via the production of a
CC vaccine or gene therapy. The current sequence is that of a Neisseria
CC meningitidis MC58 outer-membrane vesicle (OMV)-related membrane protein
CC of the invention.
XX
SQ Sequence 397 AA;
Query Match      64.1%; Score 1310; DB 8; Length 397;
Best Local Similarity 61.8%; Pred. No. 4.5e-123;
Matches 244; Conservative 64; Mismatches 87; Indels 0; Gaps 0;

QY 2 FENITAPADPILGLADLFRADERPGKINLIGVYKDTGKTPVLTTSVKAEQVLLNET 61
DB 3 FKHIAPADPILGLGEAFKAETRPKVNIGVYKDSAGATPIVKAKEAKLLSESET 62

QY 62 TKNYLIGIDGIPFGRCCTQELLFGKGSALINDKRAARTAQTPGCTGALRVAADFLAKNTSVK 121
DB 63 TKNYLTIIDGVADYNAQTQILLFGKDHEIIASRRAKTAQSLGCTGALRIAFAFAKQLNAQ 122

QY 122 RVVSNPSPWPNHKSFNVSAGLEVREYAYDAENHTLDFDALINSINEAQGDVVLFGCC 181
DB 123 TIWISNPTWPNHNAIAKAVGIQDPYRYDAAGHGLDWDGMIEDLSQAQKGDIVLLHGCC 182

QY 182 HNPTGIDTPTLEQWOTLAQLSVEKGMPLDFPAYQGFARGLEDAEGLRAPAMHKLIVA 241
DB 183 HNPTGIDTPTLEQWETLAKLSAEKGMPLDFPAYQGFNGLEBEDAYGLRVFLKHNTPELLIA 242

QY 242 SSKNFGLYNERVAGCTLVAAADSETVDRAFSPQKAAIRANYSNPPAHGASVATILSND 301
DB 243 SSKNFGLYNERVAGFTLVAADEATAARAHSQVKTIIRTLYSNPASHGANTIALVLKND 302

QY 302 ALRAIWEQELTDMRQRIQRMQLFVNTLQEKGANRDFSFIKONGMFSFGLTKEQVLR 361
DB 303 DLKAQWIAELDEMGRKAMQKQFVELLKAKGASQNFDFIIEQNGMFSFGLTPEQVDR 362

QY 362 REEFGVYVAGSRVNVAGTDPDNPAPLCEAIVAVL 396
DB 363 KNEFAIYVRSGRINVAGITDDNIDYLCESIVKVL 397

RESULT 13
ADS22791
ID ADS22791 standard; protein; 394 AA.
XX
AC ADS22791;
XX
XX ADS22791;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #11824.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
XX US2003233675-A1.
PN
XX
XX 18-DEC-2003.
PD
XX

```



```
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 396 AA;
Query Match 48.7%; Score 996; DB 8; Length 396;
Best Local Similarity 48.2%; Pred. No. 2.6e-91;
Matches 191; Conservative 67; Mismatches 138; Indels 0; Gaps 0;
QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIGVYKDETKTPTVLTSVKKAQVLLNE 60
DB 1 LFSAVELAPRDPILGLNEAFNADTRTTKVNGLGVYFNEEGKIPLLRVRAVDAEKARVDAA 60
QY 61 TTKNYLGIDGPIPEFCRGTQELLFGKGSALINDKARTATQTPGGTGALVAADFIAKNTSV 120
DB 61 LPRGYLPIEGIAAYDAAVQKLLGNDSPLIAAGRVVTTAAQALGGTGALKIGADFLLRLNPN 120
QY 121 KRVVWNSPWNHKSFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
DB 121 AKVAISDPSWENHRALFEGAGFEVVSYPYDAHTHGVNFDAMLNALSNTAAAGTIVVLHAC 180
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGMPLFPFAYQGFARGLEEDAEGRLAFAMHKEILV 240
DB 181 CHNPTGVDLNTLEQWKQVVEVKARNLVPFLDIAYQFGDNEADAAVRLFAAEELNVFV 240
QY 241 ASSYSKNFGLNERNVAGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
DB 241 SSSFSKFSGLYGERVAGLSIITASKEEAARVLSQLKRVIRTYNSNPPPHGGSVAAVLAS 300
QY 301 DALRAIWEQELTDMRQIRQRMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 360
DB 301 PELRATWETELAEMRDRIRAMRNGVLVERLKASGVDRDFSFVNAQRMFSYSGLTAPQVDR 360
QY 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
DB 361 LREEFGIYAVSTGRICVAALNRLNLDVANAIAHVL 396
RESULT 15
ADN22085
ID ADN22085 standard; protein; 398 AA.
XX
AC ADN22085;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #4738.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
```

```
XX
DR
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 4738; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition. Improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 398 AA;
Query Match 48.7%; Score 996; DB 8; Length 398;
Best Local Similarity 48.2%; Pred. No. 2.6e-91;
Matches 191; Conservative 67; Mismatches 138; Indels 0; Gaps 0;
QY 1 MFENITAAPADPILGLADLFRADERPGKINLGIGVYKDETKTPTVLTSVKKAQVLLNE 60
DB 3 LFSAVELAPRDPILGLNEAFNADTRTTKVNGLGVYFNEEGKIPLLRVRAVDAEKARVDAA 62
QY 61 TTKNYLGIDGPIPEFCRGTQELLFGKGSALINDKARTATQTPGGTGALVAADFIAKNTSV 120
DB 63 LPRGYLPIEGIAAYDAAVQKLLGNDSPLIAAGRVVTTAAQALGGTGALKIGADFLLRLNPN 122
QY 121 KRVVWNSPWNHKSFNAGLEVREYAYDAENHTLDFDALINSLNEAQAGDVVLPFGC 180
DB 123 AKVAISDPSWENHRALFEGAGFEVVSYPYDAHTHGVNFDAMLNALSNTAAAGTIVVLHAC 182
QY 181 CHNPTGIDPTLEQWOTLAQLSVEKGMPLFPFAYQGFARGLEEDAEGRLAFAMHKEILV 240
DB 183 CHNPTGVDLNTLEQWKQVVEVKARNLVPFLDIAYQFGDNEADAAVRLFAAEELNVFV 242
QY 241 ASSYSKNFGLNERNVAGACTLVAADSETVDRAFQSKAAIRANYSNPPAHGASVVATILSN 300
DB 243 SSSFSKFSGLYGERVAGLSIITASKEEAARVLSQLKRVIRTYNSNPPPHGGSVAAVLAS 302
QY 301 DALRAIWEQELTDMRQIRQRMQLFVNTLQEKGNRDFSFIIKONGMFSFSGLTKEQVLR 360
DB 303 PELRATWETELAEMRDRIRAMRNGVLVERLKASGVDRDFSFVNAQRMFSYSGLTAPQVDR 362
QY 361 LREEFGVYAVASGRVNVAGMTPDNNAPLCEAIVAVL 396
DB 363 LREEFGIYAVSTGRICVAALNRLNLDVANAIAHVL 398
Search completed: March 14, 2006, 17:56:05
Job time : 192 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2006, 02:34:21 ; Search time 1391 Seconds
(without alignments)
5706.433 Million cell updates/sec

Title: US-10-673-786A-1

Perfect score: 1191
Sequence: 1 atgttgagaacattaccgc.....cgattggcagctgctgttaa 1191

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1191	100.0	1191	9	ADB83284
2	1191	100.0	1191	13	AdS46042 Bacterial
3	1191	100.0	1191	14	AdW95403
4	1191	100.0	1331	2	AAV40259 Escherich
5	1191	100.0	1359	1	AAV71109
6	1191	100.0	14759	4	AAAS46273 DNA encod
7	1187.8	99.7	1293	1	AAV71108
8	839	70.4	1209	11	ACH99107
9	691.8	58.1	1191	10	ACF68649
10	691.8	58.1	110000	10	ACF67367_14
11	691.8	58.1	249878	10	ACF65381
12	655.8	55.1	1188	13	ADT46374 Bacterial
13	510	42.8	966	13	AdS45488 Bacterial
14	496.2	41.7	2133	5	AAAS88191
15	496.2	41.7	3222	5	AAAS90082
16	493.4	41.4	7977	13	ADT05500
17	493.4	41.4	349980	13	ADT05649 Haemophil
18	471.8	39.6	1191	6	AAAS97289
19	466.8	39.2	17381	3	AAAS1493 N. mening

C	20	466.8	39.2	110000	3	AAAS1490 05	Continuation (6 of
C	21	466.8	39.2	349980	3	AAF21607	Aaf21607 Neisseria
C	22	461.4	38.7	1269	10	ABZ40973	Abz40973 N. gonorr
C	23	458.2	38.5	110000	2	AAT42063_16	Continuation (17 o
C	24	387.6	32.5	579	11	ACH99159	ACH99159 Klebsiell
C	25	372	31.2	373	4	AAAS1443	AAAS1443 cDNA enco
C	26	363	30.5	2220	5	AAAS1804	AAAS1804 DNA encod
C	27	348.4	29.3	702	10	ADP03520	AdP03520 Bacterial
C	28	331.8	27.9	1185	13	ADSS9837	AdS9837 Bacterial
C	29	327.8	27.5	1191	13	ADSS6437	AdS6437 Bacterial
C	30	320.8	26.9	1188	13	ADSS5510	AdS5510 Bacterial
C	31	320.8	26.9	1194	13	ADSA4995	AdS4995 Bacterial
C	32	316.4	26.6	1509	11	ABD14364	ABD14364 Pseudomon
C	33	316.4	26.6	1632	11	ABD13903	ABD13903 Pseudomon
C	34	310.6	26.1	1194	13	ADSS62038	AdS62038 Bacterial
C	35	306.6	25.7	24417	2	AAT97221	AAT97221 Pseudomon
C	36	305	25.6	1242	11	ABD13824	ABD13824 Pseudomon
C	37	305	25.6	1329	11	ABD13558	ABD13558 Pseudomon
C	38	297.4	25.0	1093	5	AAAS77350	AAAS77350 DNA encod
C	39	289	24.3	1338	11	ACH95506	ACH95506 Klebsiell
C	40	288.8	24.2	1188	13	ADSS6672	AdS6672 Bacterial
C	41	287.6	24.1	1185	13	ADSS61736	AdS61736 Bacterial
C	42	285.4	24.0	1191	10	ABZ38235	Abz38235 N. gonorr
C	43	285.2	23.9	1191	13	ADSS63841	AdS63841 Bacterial
C	44	285.2	23.9	1191	13	ADSS64221	AdS64221 Bacterial
C	45	285.2	23.9	1194	13	ADSS63473	AdS63473 Bacterial

ALIGNMENTS

RESULT 1
ID ADB83284 standard; DNA; 1191 BP.
XX ADB83284;
XX 04-DEC-2003 (first entry)
XX Escherichia coli aspartate aminotransferase gene.
KW ds; gene; aspartate aminotransferase; threonine; fermentation.
XX Escherichia coli.
FH Key Location/Qualifiers
FT CDS 1..1191
FT /tag= a
FT /product= "aspartate aminotransferase"
XX WO2003072786-A1.
XX 04-SEP-2003.
XX 25-FEB-2003; 2003WO-JP002067.
XX 27-FEB-2002; 2002RU-00104983.
XX (AJIN) AJINOMOTO CO INC.
XX Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;
XX WPI; 2003-721777/68.
XX P-FSB; ADB83285.
XX Industrial production of L-threonine by fermentation using Escherichia
XX modified to enhance aspartate aminotransferase activity, with improved
XX productivity.
XX Claim 7; Page 18-20; 26pp; Japanese.
XX The invention relates to a bacterium belonging to the genus Escherichia
XX which is modified to enhance aspartate aminotransferase activity for the

CC	production of L-threonine. The bacterium is used for the industrial									
CC	production of L-threonine by fermentation. This sequence corresponds to									
XX	the E. coli aspartate aminotransferase gene.									
SQ	Sequence	1191 BP;	291 A;	305 C;	317 G;	278 T;	0 U;	0 Other;		
	Query Match	100.0%; Score 1191; DB 9; Length 1191;								
	Best Local Similarity	100.0%; Pred. No. 0;								
	Matches 1191;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1	ATGTTTGAGAACATTACCGCGCTCTCGCGACGCCGATCTTGGCGCTGGCCGATCTGTTT	60							
DB	1	ATGTTTGAGAACATTACCGCGCTCTCGCGACGCCGATCTTGGCGCTGGCCGATCTGTTT	60							
QY	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGCG	120							
DB	61	CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGCG	120							
QY	121	GGCAAAACCCGGTACTGACCGCTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA	180							
DB	121	GGCAAAACCCGGTACTGACCGCTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA	180							
QY	181	ACCACCAAAATTTACCTCGGCATTTGACGGCATCCCTGAAATTTGGTCGCTGCACTCAGGAA	240							
DB	181	ACCACCAAAATTTACCTCGGCATTTGACGGCATCCCTGAAATTTGGTCGCTGCACTCAGGAA	240							
QY	241	CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACGTCGTCGACGGCACAGACT	300							
DB	241	CTGCTGTTGGTAAAGTAGCGCCTGATCAATGACAAACGTCGTCGACGGCACAGACT	300							
QY	301	CCGGGGGCACTGGCGCACTACCGTGCGTGGCGATTTCTGCAAAATATACCGGTT	360							
DB	301	CCGGGGGCACTGGCGCACTACCGTGCGTGGCGATTTCTGCAAAATATACCGGTT	360							
QY	361	AACGCTGTGGTGAGCAACCAAGCTGGCGCAACCATGAAGCGCTTTAACTCTGCA	420							
DB	361	AACGCTGTGGTGAGCAACCAAGCTGGCGCAACCATGAAGCGCTTTAACTCTGCA	420							
QY	421	GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTTTGACTTCGAT	480							
DB	421	GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTTTGACTTCGAT	480							
QY	481	GCATGATTAACGCTGATGATGAGCTCAGGCTGGCGAGTAGTGTGTTCCATGCTGC	540							
DB	481	GCATGATTAACGCTGATGATGAGCTCAGGCTGGCGAGTAGTGTGTTCCATGCTGC	540							
QY	541	TGCCATAACCAACCGGTATCGACCCCTACGCTCGAAACATGCAACACATGGCACAACTC	600							
DB	541	TGCCATAACCAACCGGTATCGACCCCTACGCTCGAAACATGCAACACATGGCACAACTC	600							
QY	601	TCCGTTTGAGAAAGCGTGTGTTACCGCTGTTTGAATTTCGCTTACCAGGTTTGGCCGTGGT	660							
DB	601	TCCGTTTGAGAAAGCGTGTGTTACCGCTGTTTGAATTTCGCTTACCAGGTTTGGCCGTGGT	660							
QY	661	CTGGAAGAGATGCTGAGGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATGTT	720							
DB	661	CTGGAAGAGATGCTGAGGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATGTT	720							
QY	721	GCCAGTTCCTACTCTAAAACTTTGGCTCTGACAGGAGCTGTGGCGCTTGTACTCTG	780							
DB	721	GCCAGTTCCTACTCTAAAACTTTGGCTCTGACAGGAGCTGTGGCGCTTGTACTCTG	780							
QY	781	GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCG	840							
DB	781	GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCG	840							
QY	841	GCTAACTACTCTAACCCACAGACACGCGGCTTCTGTTGCGCCACCATCTCGAGCAAC	900							
DB	841	GCTAACTACTCTAACCCACAGACACGCGGCTTCTGTTGCGCCACCATCTCGAGCAAC	900							
QY	901	GATGCGTTACGTGCGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT	960							
DB	901	GATGCGTTACGTGCGATTTGGGAAACAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT	960							

QY	961	ATGCGTCAAGTTCTTCTCAATACGCTGCAGGAAAGGCGCAAAACCGGACTTCAGCTTT	1020
DB	961	ATGCGTCAAGTTCTTCTCAATACGCTGCAGGAAAGGCGCAAAACCGGACTTCAGCTTT	1020
QY	1021	ATCATCAAAACAGACGCGCATGTTCTCTTTCAGTGGCCTGACAAAAGAACAAAGTGTGCGT	1080
DB	1021	ATCATCAAAACAGACGCGCATGTTCTCTTTCAGTGGCCTGACAAAAGAACAAAGTGTGCGT	1080
QY	1081	CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGGTAAATGTGGCCGGATG	1140
DB	1081	CTGCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGGTAAATGTGGCCGGATG	1140
QY	1141	ACACAGATAAATCATGCTCCGCTGTCGAAGCATTTGTGGCAGTGTCTGTA	1191
DB	1141	ACACAGATAAATCATGCTCCGCTGTCGAAGCATTTGTGGCAGTGTCTGTA	1191

RESULT 2

ADS46042	standard; cDNA; 1191 BP.
AC	ADS46042;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Bacterial polynucleotide #785.
XX	
KW	Recombinant DNA construct; transformed plant; improved plant property;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW	pathogen tolerance; pest tolerance; plant disease resistance;
KW	cell cycle pathway modification; plant growth regulator;
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX	bacterial polynucleotide; gene; ss.
OS	Bacteria.
XX	
PN	US2003233675-A1.
XX	
PD	18-DEC-2003.
XX	
PF	20-FEB-2003; 2003US-00369493.
XX	
PR	21-FEB-2002; 2002US-0360039P.
XX	
PA	(CAOY/) CAO Y.
PA	(HINK/) HINKLE G J.
PA	(SLAT/) SLATER S C.
PA	(CHEN/) CHEN X.
PA	(GOLD/) GOLDMAN B S.
XX	
PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	
DR	WPI; 2004-061375/06.
XX	
PT	New recombinant DNA construct comprising a promoter positioned to provide
PT	for expression of a polynucleotide encoding a polypeptide from a
PT	microbial source, useful for producing plants with improved properties.
XX	
PS	Claim 1; SEQ ID NO 24472; 122pp; English.
XX	
CC	The invention relates to a recombinant DNA construct comprising a
CC	promoter functional in a plant cell, where the promoter is positioned to
CC	provide for expression of a polynucleotide encoding a polypeptide from a
CC	microbial source. The invention also relates to a transformed plant
CC	comprising the recombinant DNA construct and a method of producing a
CC	transformed plant having an improved property. The plant is a crop plant
CC	such as maize or soybean. The method of producing a transformed plant
CC	having an improved property comprises transforming a plant with the
CC	recombinant DNA construct and growing the transformed plant, where the
CC	polynucleotide or polypeptide is useful for improving plant properties.
CC	The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

Query Match 100.0%; Score 1191; DB 13; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCGATTCTGGGCTTGGCGGATCTGTTT 60
DB 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCCGATTCTGGGCTTGGCGGATCTGTTT 60
QY 61 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG 120
DB 61 CGTGGCGATGAACGTCCTCGGCAAAATTAACCTCGGATTGGTGTCTATAAAGATGAGACG 120
QY 121 GGCAAAACCCGGTACTACGACGGTGAAGAGGCTGAACAGTATCTGCTCGAAAATGAA 180
DB 121 GGCAAAACCCGGTACTACGACGGTGAAGAGGCTGAACAGTATCTGCTCGAAAATGAA 180
QY 181 ACCACCAAAATTAACCTCGGCAATTAACCTCGGATTCCTGCACTCAGGAA 240
DB 181 ACCACCAAAATTAACCTCGGCAATTAACCTCGGATTCCTGCACTCAGGAA 240
QY 241 CTGCTGTTGGTAAAGTACGCGCTGATCAATGACAAACGTCTCGCAGGACAGACT 300
DB 241 CTGCTGTTGGTAAAGTACGCGCTGATCAATGACAAACGTCTCGCAGGACAGACT 300
QY 301 CCGGGGGCACTGGCGCACTACCGGTGGCTGCGGATTTCTGGCAAAAATACACGGTT 360
DB 301 CCGGGGGCACTGGCGCACTACCGGTGGCTGCGGATTTCTGGCAAAAATACACGGTT 360
QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGCCCAACCAATGAAGCGTCTTAACTCTGCA 420
DB 361 AAGCGTGTGGGTGAGCAACCAAGCTGCCCAACCAATGAAGCGTCTTAACTCTGCA 420
QY 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
DB 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
QY 481 GCACGTATTAACAGCTGAATGAAGCTCAGGCTGGCGACGTAAGTGTGTTCCATGCTGC 540
DB 481 GCACGTATTAACAGCTGAATGAAGCTCAGGCTGGCGACGTAAGTGTGTTCCATGCTGC 540
QY 541 TGCCATAACCAACCGGTATCGACCTACGCTGGAACCAATGGCAACACTGGCACAACTC 600
DB 541 TGCCATAACCAACCGGTATCGACCTACGCTGGAACCAATGGCAACACTGGCACAACTC 600
QY 601 TCGTTTGAGAAAGGCTGTTACCGCTGTTTGAATCTCGCTTACCAGGGTTTGGCCGCTGGT 660
DB 601 TCGTTTGAGAAAGGCTGTTACCGCTGTTTGAATCTCGCTTACCAGGGTTTGGCCGCTGGT 660
QY 661 CTGGAAGAAGAGTCTGAAGAGACTGCGCGCTTTTCGGGCTATGATGAAGAGCTGATGTT 720
DB 661 CTGGAAGAAGAGTCTGAAGAGACTGCGCGCTTTTCGGGCTATGATGAAGAGCTGATGTT 720
QY 721 GCGAGTTCCTACTCTAAAACCTTGGCTGTCAACAGCGGTGTTGGCGCTTGACTCTG 780
DB 721 GCGAGTTCCTACTCTAAAACCTTGGCTGTCAACAGCGGTGTTGGCGCTTGACTCTG 780
QY 781 GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGGCCAATGAAGAGCGGATTCG 840

DB 781 GTTGCTGCCGACAGTGAACCCGTTGATCGCGCATTCAGCCAAATGAAGAGCGGATTCGC 840
QY 841 GCTAACTACTCTAACCCACCAGCACACGGCGCTTCTGTTGTCACCATCTCCTGAGCAAC 900
DB 841 GCTAACTACTCTAACCCACCAGCACACGGCGCTTCTGTTGTCACCATCTCCTGAGCAAC 900
QY 901 GATCGCTTACGTCGATTTTCGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
DB 901 GATCGCTTACGTCGATTTTCGGAAACAAGAGCTGACTGATATGCGCCAGCGTATTTCAGCGT 960
QY 961 ATGCGTCAGTGTTCGTCATACGTCGACGAAAAAGCGCAAAACCCGCACTTCAGCTTT 1020
DB 961 ATGCGTCAGTGTTCGTCATACGTCGACGAAAAAGCGCAAAACCCGCACTTCAGCTTT 1020
QY 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAGTGCCTCGT 1080
DB 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAGTGCCTCGT 1080
QY 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTCGCGTAAATGTGGCCGGATG 1140
DB 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTCGCGTAAATGTGGCCGGATG 1140
QY 1141 ACACGAGATAACATGCTCGCTGTCGAAAGCGATTTGTCGAGTGTGTA 1191
DB 1141 ACACGAGATAACATGCTCGCTGTCGAAAGCGATTTGTCGAGTGTGTA 1191
RESULT 3
ADM95403
ID ADM95403 standard; DNA; 1191 BP.
XX
AC ADM95403;
XX
DT 07-APR-2005 (first entry)
XX
DE Nucleotide sequence of an Escherichia coli polypeptide.
XX
KW transgenic plant; glutamic acid dehydrogenase; GDH; ECASPC;
KW 2-oxo glutaric acid; agriculture; gene; ds.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..1191
FT /*tag= a
XX
PN WO2005006847-A1.
XX
PD 27-JAN-2005.
XX
PF 15-JUL-2004; 2004WO-JP010451.
XX
PR 17-JUL-2003; 2003JP-00198559.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Kisaka H, Miwa T, Akiyama A;
XX
DR WPI; 2005-132242/14.
XX
PT P-PSDB; ADM95404.
XX
PT Producing plant having improved growth and yield under cultivation
PT conditions with decrease in nitrogen, involves expressing transduced
PT glutamic acid dehydrogenase gene in plant so as to increase their 2-oxo
PT glutaric acid contents.
XX
PS Disclosure; SEQ ID NO 19; 94pp; Japanese.
XX
CC The specification describes a method of producing a plant having improved
CC growth and yield under cultivation conditions with a decrease in
CC nitrogen. The method involves transducing glutamic acid dehydrogenase
CC (GDH) or ECASPC genes into a plant and expressing the gene so as to

CC increase 2-oxo glutaric acid content of the plant, or applying proline to
CC the foliage of the plant so as to increase the 2-oxo glutaric acid
CC content of the plant. The method is useful for producing a plant having
CC improved growth and yield under cultivation conditions with a decrease in
CC nitrogen. The present sequence encodes a polypeptide, used in the course
CC of the invention.
XX
SQ Sequence 1191 BP; 291 A; 305 C; 317 G; 278 T; 0 U; 0 Other;

Query Match 100.0%; Score 1191; DB 14; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAGAACTTACCGCGCTCTCGCGACCCCGAATCTGGGCCCTGGCCGATCTGTTT 60
DB 1 ATGTTTGAGAACTTACCGCGCTCTCGCGACCCCGAATCTGGGCCCTGGCCGATCTGTTT 60
QY 61 CGTGGCGATGACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGCG 120
DB 61 CGTGGCGATGACGTCCTCCGCAAAATTAACCTCGGATTTGCTCTATAAGATGAGCG 120
QY 121 GGCAAAACCCCGGTACTACCGCGTGAAAGAGCTGAACAGTATCTGCTCGAAAATGAA 180
DB 121 GGCAAAACCCCGGTACTACCGCGTGAAAGAGCTGAACAGTATCTGCTCGAAAATGAA 180
QY 181 ACCACAAAATTAACCTCGGATTTGACGGGATCCCTGAAATTTGGTGGCTGCACTCAGGAA 240
DB 181 ACCACAAAATTAACCTCGGATTTGACGGGATCCCTGAAATTTGGTGGCTGCACTCAGGAA 240
QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGACAGACT 300
DB 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCAGGACAGACT 300
QY 301 CGGGGGGCACTGGCGCACTACCGTGCGTGGCGATTTCTTGCAAAAAATACACGCTT 360
DB 301 CGGGGGGCACTGGCGCACTACCGTGCGTGGCGATTTCTTGCAAAAAATACACGCTT 360
QY 361 AAGCGTGTGGGTAGCAACCCAGCTGCGCGAACCATAAGAGCGTCTTTAACTCTGCA 420
DB 361 AAGCGTGTGGGTAGCAACCCAGCTGCGCGAACCATAAGAGCGTCTTTAACTCTGCA 420
QY 421 GGTCTGGAGTTCGTGAATACGCTTATATGATGCGGAAATCACAACCTTCACTTCGAT 480
DB 421 GGTCTGGAGTTCGTGAATACGCTTATATGATGCGGAAATCACAACCTTCACTTCGAT 480
QY 481 GCACTGATTAAACAGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCATGCTGC 540
DB 481 GCACTGATTAAACAGCTGAATGAAGCTCAGGCTGGCGAGCTAGTGTGTTCATGCTGC 540
QY 541 TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACATGGCACAATC 600
DB 541 TGCCATAACCCAAACCGGTATCGACCCCTACGCTGGAACAATGGCAAAACACATGGCACAATC 600
QY 601 TCCGTTTGAGAAAGCGTGTACCGCTGTTGACTTCGCTTACCAGGTTTTCGCCGTGT 660
DB 601 TCCGTTTGAGAAAGCGTGTACCGCTGTTGACTTCGCTTACCAGGTTTTCGCCGTGT 660
QY 661 CTGGAAGAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATGAAGAGCTGATTGTT 720
DB 661 CTGGAAGAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATGAAGAGCTGATTGTT 720
QY 721 GCCAGTTCCTACTCTAATAAATTTGGCTGTACAAGAGCGTGTGGCGCTTTGACTCTG 780
DB 721 GCCAGTTCCTACTCTAATAAATTTGGCTGTACAAGAGCGTGTGGCGCTTTGACTCTG 780
QY 781 GTTGCTGCCGACAGTGAACCGTGTATCGCGATTCAGCCAAATGAAGCGGCGATTCCG 840
DB 781 GTTGCTGCCGACAGTGAACCGTGTATCGCGATTCAGCCAAATGAAGCGGCGATTCCG 840
QY 841 GCTAACTACTCTAACCCACGACACAGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900
DB 841 GCTAACTACTCTAACCCACGACACAGCGGCTTCTGTTGTTGCCACCATCTCTGAGCAAC 900

QY 901 GATGCGTTACGTCGATTTGGGAAACAAGAGCTGACTGATATGCGGCAGCGTATTACGCGT 960
DB 901 GATGCGTTACGTCGATTTGGGAAACAAGAGCTGACTGATATGCGGCAGCGTATTACGCGT 960
QY 961 ATGCGTCAAGTTGTTTCGTCGAATACGCTGAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
DB 961 ATGCGTCAAGTTGTTTCGTCGAATACGCTGAGGAAAAAGCGCAAAACCGGACTTCAGCTTT 1020
QY 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCCTGACAAAAGAACAGTGTGCGCT 1080
DB 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCCTGACAAAAGAACAGTGTGCGCT 1080
QY 1081 CTGCGCGAAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTGCCCGGATG 1140
DB 1081 CTGCGCGAAGAGTTTGGCGTATATCGGTTGCTTCTGTCGCGTAAATGTGCCCGGATG 1140
QY 1141 ACACAGAGATAACATGGCTCCGCTGTCGAAGCGATTGTGGCAGTGTCTGTA 1191
DB 1141 ACACAGAGATAACATGGCTCCGCTGTCGAAGCGATTGTGGCAGTGTCTGTA 1191

RESULT 4
AAV40259

ID AAV40259 standard; DNA; 1331 BP.

AC AAV40259;

XX 13-OCT-1998 (first entry)

XX Escherichia coli aspC gene.

XX Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;

KW aspartokinase; feedback inhibition; dihydropicolinate reductase;

KW diaminopimelate decarboxylase; aspartate aminotransferase; ds.

XX Escherichia coli.

Key Location/Qualifiers

FT 10..1197

FT /*tag= a

FT /product= "aspC"

XX EP854189-A2.

XX 22-JUL-1998.

XX 05-DEC-1997; 97EP-00121443.

XX 05-DEC-1996; 96JP-00325659.

XX (AJIN) AJINOMOTO CO INC.

XX Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;

XX WPI; 1998-379060/33.

XX P-PSDB; AAW69553.

XX Recombinant DNA autonomously replicable in coryneform bacteria - used to
PT produce L-lysine, codes for e.g. aspartokinase, dihydropicolinate
PT reductase and synthase and di:amino-pimelate decarboxylase.

XX Example 5; Page 37-38; 59pp; English.

XX The present invention describes a recombinant DNA autonomously replicable
CC in cells of coryneform bacteria (CB), comprising a DNA sequence coding
CC for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
CC threonine is desensitised, a DNA sequence coding for a
CC dihydropicolinate reductase (DHRP), a DNA sequence coding for
CC dihydropicolinate synthase (DHPS), a DNA sequence coding for
CC diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
CC aspartate aminotransferase (AAT). The present sequence encodes aspC from
CC Escherichia coli. The DNA and related products from the present
CC invention, can be used for improving L-lysine productivity by CB. The L-

CC lysine produced can be used as a fodder additive

XX SQ Sequence 1331 BP; 330 A; 340 C; 350 G; 311 T; 0 U; 0 Other;
Query Match 100.0%; Score 1191; DB 2; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTTCAGAACATTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGGCGGATCTGTTT 60
DB 10 ATGTTTTCAGAACATTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGGCGGATCTGTTT 69
QY 61 CGTGCCGATGAACGTCCTCGGCAAAATTAACCTCGGATTCGTTCTATAAAGATGAGACG 120
DB 70 CGTGCCGATGAACGTCCTCGGCAAAATTAACCTCGGATTCGTTCTATAAAGATGAGACG 129
QY 121 GGCNAACCCCGTACTGACGAGCGTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA 180
DB 130 GGCNAACCCCGTACTGACGAGCGTGAAAGAGCTGAACAGTATCTGCTCGAAATGAA 189
QY 181 ACCACCAAAATTAACCTCGGCAATTCGCGCATCCCTGAATTTGGTCTGCTGCACCTCAGGAA 240
DB 190 ACCACCAAAATTAACCTCGGCAATTCGCGCATCCCTGAATTTGGTCTGCTGCACCTCAGGAA 249
QY 241 CTGCTGTTTGGTAAAGTAGCGCTCTGATCAATGACAAACGTCGTCGACGCGCACAGCT 300
DB 250 CTGCTGTTTGGTAAAGTAGCGCTCTGATCAATGACAAACGTCGTCGACGCGCACAGCT 309
QY 301 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGGATTTCTTGGCAAAAAATACGAGGTT 360
DB 310 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGGATTTCTTGGCAAAAAATACGAGGTT 369
QY 361 AAGCGTGTGGGTGAGCAACCAAGCTGCGCAACCAAGCTGCGCAACCAAGCTGCGCAACCAAGCTGCGCA 420
DB 370 AAGCGTGTGGGTGAGCAACCAAGCTGCGCAACCAAGCTGCGCAACCAAGCTGCGCAACCAAGCTGCGCA 429
QY 421 GGTCTGGAAGTTCGTGTAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 480
DB 430 GGTCTGGAAGTTCGTGTAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 489
QY 481 GCACCTGTAATACAGCTGTAATGAGCTCAGGCTGGCGACGTAAGTGTGTTTCCATGGCTGC 540
DB 490 GCACCTGTAATACAGCTGTAATGAGCTCAGGCTGGCGACGTAAGTGTGTTTCCATGGCTGC 549
QY 541 TGCCTAATACCAACCGGTATCGACCTTACCTGGAACAATGGCAACCACTGGCACAACTC 600
DB 550 TGCCTAATACCAACCGGTATCGACCTTACCTGGAACAATGGCAACCACTGGCACAACTC 609
QY 601 TCGGTTGAGAAAGCGTGTACCGCTGTTTGACTTTCGCTTACGAGGTTTTCGCCGTGGT 660
DB 610 TCGGTTGAGAAAGCGTGTACCGCTGTTTGACTTTCGCTTACGAGGTTTTCGCCGTGGT 669
QY 661 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATGTTT 720
DB 670 CTGGAAGAAGATGCTGAAGACTGCGCGCTTTTCGCGCTATGATGAAGAGCTGATGTTT 729
QY 721 GCCAGTCTCTACTATAAAATTTGGCTGTGACAGAGCGTGTGGCGCTTGTACTCTG 780
DB 730 GCCAGTCTCTACTATAAAATTTGGCTGTGACAGAGCGTGTGGCGCTTGTACTCTG 789
QY 781 GTTGCTGCCGACAGTGAACCGTTGATGCGGCTTTCAGCCAAATGAAGCGGCGATTCGC 840
DB 790 GTTGCTGCCGACAGTGAACCGTTGATGCGGCTTTCAGCCAAATGAAGCGGCGATTCGC 849
QY 841 GCTAACTACTCTAACCCACGACACGCGGCTTCTGTTGTGCCACCATCTCTGAGCAAC 900
DB 850 GCTAACTACTCTAACCCACGACACGCGGCTTCTGTTGTGCCACCATCTCTGAGCAAC 909
QY 901 GATGCGTTAGCTGGATTTGGGAACAAGACTGACTGATATGCGCCAGCGTATTCAGCGT 960
DB 910 GATGCGTTAGCTGGATTTGGGAACAAGACTGACTGATATGCGCCAGCGTATTCAGCGT 969
QY 961 ATGCGTCAGTTGTTTCGTCATACGCTGCGAGGAAAGCGCGCAAAACCGCGACTTCAGCTTT 1020

DB 970 ATGCGTCAGTTGTTTCGTCATACGCTGACGAAAAAGCGCAAAACCGCACTTCAGCTTT 1029
QY 1021 ATCATCAAAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAGTCTGCGT 1080
DB 1030 ATCATCAAAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAGTCTGCGT 1089
QY 1081 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGCGCGGATG 1140
DB 1090 CTGCGCAAGAGTTTGGCGTATATGCGGTTGCTTCTGCTGCGTAAATGTGCGCGGATG 1149
QY 1141 ACACCAAGATTAACATGGCTCCGCTGTGCGAAGCGATTTGCGAGTGTGTTAA 1191
DB 1150 ACACCAAGATTAACATGGCTCCGCTGTGCGAAGCGATTTGCGAGTGTGTTAA 1200

RESULT 5

AAN71109

ID AAN71109 standard; DNA; 3659 BP.

XX AC AAN71109;

XX AC AAN71109;

DT 01-JAN-1980 (first entry)

XX pheA aroF aspC operon in plasmid pME219.

XX alpha-amylase; feedback inhibition; amino acid synthesis;

XX composite plasmid; ss.

XX Bacillus licheniformis.

Key Location/Qualifiers

FT CDS 69..1241

FT /*tag= a

FT CDS 1259..2329

FT /*tag= b

FT CDS 2344..3534

FT /*tag= c

XX WO800202-A.

XX 15-JAN-1987.

XX 24-JUN-1986; 86WO-US001353.

XX 24-JUN-1985; 85US-00747732.

XX (NUTR-) NUTRASWEET CO.

XX Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;

XX WPI; 1987-021998/03.

DR P-PSDB; AAP70752, AAP71677, AAP71678.

XX Composite plasmids contg. multiple genes in transcriptional units -

PT useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.

XX Disclosure; Page 38; 57pp; English.

XX This sequence may be inserted into a composite plasmid and used for the

CC production of amino acids. See also AAN71053-55, AAN71107, AAN71109,

CC AAN71111 and AAP70696-97 and AAP70750, AAP70752 and AAP70754

XX Sequence 3659 BP; 936 A; 911 C; 947 G; 865 T; 0 U; 0 Other;

Query Match 100.0%; Score 1191; DB 1; Length 3659;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTTCAGAACATTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGGCGGATCTGTTT 60

DB 2344 ATGTTTTCAGAACATTACCGCGCTCTCGCGACCCGATTCCTGGCCCTGGCGGATCTGTTT 2403

QY 61 CGTCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG 120
DB 2404 CGTGCCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG 2463
QY 121 GGCAAAACCCCGGTACTACGACGCTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA 180
DB 2464 GGCAAAACCCCGGTACTACGACGCTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA 2523
QY 181 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTAGGAA 240
DB 2524 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTAGGAA 2583
QY 241 CTGCTGTTGGTAAAGGTAGCGCTGATCAATGACAAACGCTGCGACGGCACAGACT 300
DB 2584 CTGCTGTTGGTAAAGGTAGCGCTGATCAATGACAAACGCTGCGACGGCACAGACT 2643
QY 301 CCGGGGGGCACTGGCGCACTACCGCTGGCTGCGATTTCTGGCAAAATAACACGCTT 360
DB 2644 CCGGGGGGCACTGGCGCACTACCGCTGGCTGCGATTTCTGGCAAAATAACACGCTT 2703
QY 361 AAGCGTGTGGTGAAGCAACCCAGCTGGCGCAACCAATGAAGCGCTTTAACTCTGCA 420
DB 2704 AAGCGTGTGGTGAAGCAACCCAGCTGGCGCAACCAATGAAGCGCTTTAACTCTGCA 2763
QY 421 GGTCTGGAAGTTCTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 480
DB 2764 GGTCTGGAAGTTCTGAATACGCTTATATGATGCGGAAATCACACTCTTGACTTCGAT 2823
QY 481 GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTCTTCCATGCTGC 540
DB 2824 GCACTGATTAACAGCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTCTTCCATGCTGC 2883
QY 541 TGCATTAACCAACCCGATTCGACCTACGCTGGAACAATGGCAAACTGGGCACAATC 600
DB 2884 TGCATTAACCAACCCGATTCGACCTACGCTGGAACAATGGCAAACTGGGCACAATC 2943
QY 601 TCGTTGAGAAAGCTGTTACCGCTTTTGACTTCGCTTACCAGGTTTTCGCCGTGCT 660
DB 2944 TCGTTGAGAAAGCTGTTACCGCTTTTGACTTCGCTTACCAGGTTTTCGCCGTGCT 3003
QY 661 CTGGAAGAGATGCTCAAGGACTCGCGCTTTTCGCGCTATGATAAAGAGCTGATGTT 720
DB 3004 CTGGAAGAGATGCTCAAGGACTCGCGCTTTTCGCGCTATGATAAAGAGCTGATGTT 3063
QY 721 GCCAGTTCCTACTCTAAAACTTTGGCTGTACAACGAGCGTTGGCGCTTGTACTCTG 780
DB 3064 GCCAGTTCCTACTCTAAAACTTTGGCTGTACAACGAGCGTTGGCGCTTGTACTCTG 3123
QY 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCCG 840
DB 3124 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGATTCCG 3183
QY 841 GCTAACTACTCTAACCCACGACACGCGCTTCTGTTGTCGCCACCATCTTGAGCAAC 900
DB 3184 GCTAACTACTCTAACCCACGACACGCGCTTCTGTTGTCGCCACCATCTTGAGCAAC 3243
QY 901 GATGCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
DB 3244 GATGCGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 3303
QY 961 ATGCGTCAGTTGTTGTCATACGCTGCAGGAAAGGCGCAACCCGCGACTTCAGCTTT 1020
DB 3304 ATGCGTCAGTTGTTGTCATACGCTGCAGGAAAGGCGCAACCCGCGACTTCAGCTTT 3363
QY 1021 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTGGCTGCGT 1080
DB 3364 ATCATCAACAGAACGCGATGTTCTCTTCAGTGGCTGACAAAGAACAAAGTGGCTGCGT 3423
QY 1081 CTCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTCGCCGGATG 1140
DB 3424 CTCGCGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTCGCGTAAATGTCGCCGGATG 3483
QY 1141 ACACGAGATAACATGGCTCGCTGTGGGAAGCGATTGTGGCAGTGTCTGTA 1191

DB 3484 ACACGAGATAACATGGCTCGCTGCGAAGCGATTGTGGCAGTGTCTGTA 3534
RESULT 6
AAS46273/c
ID AAS46273 standard; DNA; 14759 BP.
XX AAS46273;
XX 18-DEC-2001 (first entry)
XX DNA encoding novel mar regulated protein (NIMR) #42.
XX mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.
XX Escherichia coli.
XX WO200170776-A2.
XX 27-SEP-2001.
XX 08-MAR-2001; 2001WO-US007478.
XX 10-MAR-2000; 2000US-0188362P.
XX (TUFT) TUFTS COLLEGE.
XX Levy SB, Barbosa TM, Alekshun MN;
XX WPI: 2001-602769/68.
XX P-PSDB; AAU29374.
XX Identifying compounds that modulate a newly identified mar regulated polypeptide activity, useful as antimicrobial compounds, involves contacting the polypeptide with a test compound.
XX Disclosure; Page 477-485; 526pp; English.
XX The invention relates to a method of identifying compounds that modulate a newly identified mar regulated (NIMR) polypeptide activity. The method comprises contacting an NIMR polypeptide with a test compound under interaction conditions, determining the ability of the compound to modulate the activity or expression of the polypeptide, and selecting the modulators. NIMR nucleic acids and polypeptides are used in the treatment of microbial infections, and in screening for modulators of NIMR expression and activity. These modulators can be used to reduce the infectivity of a microbe on a surface, and the virulence of a microbe in a subject suffering from an infection. AAS46232-AAS46278 represent Escherichia coli NIMR coding sequences of the invention
XX Sequence 14759 BP; 3703 A; 3840 C; 3713 G; 3503 T; 0 U; 0 Other;
Query Match 100.0%; Score 1191; DB 4; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGAACATTAACCGCCGCTCTGCGACCCGATTTCTGGCCCTGCGCCGATCTGTTT 60
DB 2755 ATGTTTGAACATTAACCGCCGCTCTGCGACCCGATTTCTGGCCCTGCGCCGATCTGTTT 2696
QY 61 CGTGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG 120
DB 2695 CGTGCGATGAACGTCCTCCGCAAAATTAACCTCGGATTTGGTCTCTATAAAGATGAGACG 2636
QY 121 GGCAAAACCCCGGTACTACGACGCTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA 180
DB 2635 GGCAAAACCCCGGTACTACGACGCTGAAAGAGCTGAAACAGTATCTGCTCGAAATGAA 2576
QY 181 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTAGGAA 240
DB 2575 ACCACCAAAATTAACCTCGGATTTGACGCGATCCCTGAAATTTGGTCTGCTGCACTAGGAA 2516

Qy 241 CTGCTGTTTGTAAAGGTAGCGCCTGATCAATGACAAACGTGTCGACGGCAGACT 300
Db 2515 CTGCTGTTTGTAAAGGTAGCGCCTGATCAATGACAAACGTGTCGACGGCAGACT 2456
Qy 301 CCGGGGGCACTGGGGCACTAGCGTGGCTGCCGATTTCCCTGGCAAAAATACCAGGTT 360
Db 2455 CCGGGGGCACTGGGGCACTAGCGTGGCTGCCGATTTCCCTGGCAAAAATACCAGGTT 2396
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 2395 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 2336
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAATACACTCTTTGACTTCGAT 480
Db 2335 GGTCTGGAAGTTCGTGAATACGCTTATATGATCGGAAATACACTCTTTGACTTCGAT 2276
Qy 481 GCACGTATTACAGCCTGAATGAAGCTCAGCTGGCGACGTAGTGTCTTCATGGCTGC 540
Db 2275 GCACGTATTACAGCCTGAATGAAGCTCAGCTGGCGACGTAGTGTCTTCATGGCTGC 2216
Qy 541 TGCATTAACCCAAACCGGTATCGACCTACGCTGGAACAATGGCAAAACACTGGCAAACTC 600
Db 2215 TGCATTAACCCAAACCGGTATCGACCTACGCTGGAACAATGGCAAAACACTGGCAAACTC 2156
Qy 601 TCCGTTGAGAAAGCTGGTTACCGCTGTTTACCTTTCGCTTACCAGGGTTTTGCCCGTGGT 660
Db 2155 TCCGTTGAGAAAGCTGGTTACCGCTGTTTACCTTTCGCTTACCAGGGTTTTGCCCGTGGT 2096
Qy 661 CTGGAAGAAGTGTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAGAGCTGATGTT 720
Db 2095 CTGGAAGAAGTGTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAGAGCTGATGTT 2036
Qy 721 GCCAGTTCCTACTCTAAACCTTTGGCTGTGACACGAGCGTGTGGCGCTTGACTCTG 780
Db 2035 GCCAGTTCCTACTCTAAACCTTTGGCTGTGACACGAGCGTGTGGCGCTTGACTCTG 1976
Qy 781 GTTGCTGCCGACAGTAAACCGTTGATCGCGCATTTACGCCAAATGAAAGCGCGATTCGC 840
Db 1975 GTTGCTGCCGACAGTAAACCGTTGATCGCGCATTTACGCCAAATGAAAGCGCGATTCGC 1916
Qy 841 GCTAACTACTCTAAACCCACGAGCACCGCGCTTTCTGTTGTTGCCACCATCTGAGCAAC 900
Db 1915 GCTAACTACTCTAAACCCACGAGCACCGCGCTTTCTGTTGTTGCCACCATCTGAGCAAC 1856
Qy 901 GATCGGTACGTGCGATTTGGAAACAAGAGCTGACTGATATGCGCGAGCGTATTCAGGT 960
Db 1855 GATCGGTACGTGCGATTTGGAAACAAGAGCTGACTGATATGCGCGAGCGTATTCAGGT 1796
Qy 961 ATGCGTCACTGTTTCGTCAATACGCTGCAGGAAAGCGCAACCGCGACTTCAGCTTT 1020
Db 1795 ATGCGTCACTGTTTCGTCAATACGCTGCAGGAAAGCGCAACCGCGACTTCAGCTTT 1736
Qy 1021 ATCATCAACAGAAAGCGCATGTTCTCTTTCAGTGGCCTGCAAAAGAACAAAGTCTCGT 1080
Db 1735 ATCATCAACAGAAAGCGCATGTTCTCTTTCAGTGGCCTGCAAAAGAACAAAGTCTCGT 1676
Qy 1081 CTGCGCAAGATTTGGCGTATATGCGGTTCGCTTCTGCTGCGTAAATGTGGCCGGATG 1140
Db 1675 CTGCGCAAGATTTGGCGTATATGCGGTTCGCTTCTGCTGCGTAAATGTGGCCGGATG 1616
Qy 1141 ACACAGATACATGGTTCGCTGCGAAGCGATTTGGCGAGTGTGTA 1191
Db 1615 ACACAGATACATGGTTCGCTGCGAAGCGATTTGGCGAGTGTGTA 1565

RESULT 7
ID AAN71108
XX AAN71108 standard; DNA; 1293 BP.
AC AAN71108;
XX AAN71108;
DT 01-JAN-1980 (first entry)
XX

Optimized Escherichia coli aspC gene.
aspC gene; feedback inhibition; amino acid synthesis; composite plasmid;
88.
Escherichia coli.
W08700202-A.
15-JAN-1987.
24-JUN-1986; 86WO-US001353.
24-JUN-1985; 85US-00747732.
(NUTR-) NUTRASWEET CO.
Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;
WPI; 1987-021998/03.
P-PSDB; AAP70751.
Composite plasmids contg. multiple genes in transcriptional units -
useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.
Disclosure; Page 25; 57pp; English.
This sequence may be inserted into a composite plasmid and used for the
production of amino acids. See also AAN71053-55, AAN71107, AAN71109-11
and AAP70696-97 and AAP70750, AAP70752-54
SQ Sequence 1293 BP; 322 A; 331 C; 343 G; 297 T; 0 U; 0 Other;
Query Match 99.7%; Score 1187.8; DB 1; Length 1293;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGTTTTCAGAACATTACCGCGCTCTCGCGACCCGATTTGGGCCCTGGCCGATCTGTTT 60
Db 12 ATGTTTTCAGAACATTACCGCGCTCTCGCGACCCGATTTGGGCCCTGGCCGATCTGTTT 71
Qy 61 CGTGCCGATGAAGCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 120
Db 72 CGTGCCGATGAAGCTCCCGGCAAAATTAACCTCGGATTTGGTGTCTATAAAGATGAGACG 131
Qy 121 GGCACAAACCCCGGTACTGACCCAGCGTGAAAAGCTGAAAGTATCTGCTCGAAAAATGAA 180
Db 132 GGCACAAACCCCGGTACTGACCCAGCGTGAAAAGCTGAAAGTATCTGCTCGAAAAATGAA 191
Qy 181 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAATTTGGTCCGTCACCTCAGGAA 240
Db 192 ACCACCAAAATTTACCTCGGCATTGACGGCATCCCTGAATTTGGTCCGTCACCTCAGGAA 251
Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGACAAACGTGCTCGCACGGCAGACT 300
Db 252 CTGCTGTTTGGTAAAGGTAGCGCCTGATCAATGACAAACGTGCTCGCACGGCAGACT 311
Qy 301 CCGGGGGCACTGGCGCACTACGCGTGGCTGGCGATTTCTTGGCAAAAATACCAGCGTT 360
Db 312 CCGGGGGCACTGGCGCACTACGCGTGGCTGGCGATTTCTTGGCAAAAATACCAGCGTT 371
Qy 361 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 372 AAGCGTGTGGGTGAGCAACCCAAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 431
Qy 421 GGTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAATACACTCTTTGACTTCGAT 480
Db 432 GGTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAATACACTCTTTGACTTCGAT 491
Qy 481 GCACTGATTAAAGCCTGAATGAAGCTCAGCTGGCGACGTAGTGTCTTCCATGGCTGC 540
Db 492 GCACTGATTAAAGCCTGAATGAAGCTCAGCTGGCGACGTAGTGTCTTCCATGGCTGC 551

Qy 781 GTTCTGCGCAGTGAACCGTTGATCGCGCATTCAGCCAAATGAAGCGCGGATTCGC 840
 Db 799 GTCCCGCGGATCAGGAGACTGTAGACCGGCGCTTCAGTCAGATGAAGTGGTGCATCCG 858
 Qy 841 GCTAACTACTTAACCCACAGCAGACAGCGCGCTTCTGTTGTTGCCACCATCTGAGCAAC 900
 Db 859 GCCAACTACTCGAACC CGCGCTGCGCATGGCGCCTCCGTTGTCGCCACCATTCAGCAAC 918
 Qy 901 GATCGGTTACGTCGATTTGGGAACAAGAGCTGACTGATATGCGCCAGCGTATTACAGGT 960
 Db 919 GATCGGTCAGGGAATCTGGAGAGGAAGTACCGATATGCGCCAGCGCATCCAGGT 978
 Qy 961 ATCGCTCAGTTGTTGCTCAATACGCTGCAGAAAAAGCGCAACACCGCACTTCAGCTTT 1020
 Db 979 ATCGCTCTGCTGTTGCTCAATACCTCGCAGGAGAAAGCGCGAGCGCGACTTCAGCTTT 1038
 Qy 1021 ATCATCAACAGAACGCGATGTTCTCTTCTAGTGGCTGACAAAGAACAAAGTCTCGGT 1080
 Db 1039 ATCAGCAGCAGAAACGCGCATGTTCTCATTCAGCGGCGCTGACTAAAGAGCAGGTGCTCGC 1098
 Qy 1081 CTGCGCAAGAGTTTGGCGTATATGCGTGTCTTCTGTCGCGTAAATGTGGCGCGGATG 1140
 Db 1099 CTGCGTGAAGATTGCGCATCTATGCGTAGCTTCCGAGCATATCAACGTTGCGCGGATG 1158
 Qy 1141 ACACAGATAACATGGCTCCGCTGTGGGAAGCGATTGTGGCAGTGTCTGTAA 1191
 Db 1159 ACGCTGACATATGGCGCGCTGTGGAGCCATCTGTCGCGTACTGTAA 1209

RESULT 9

ACF68649
 ID ACF68649 standard; DNA; 1191 BP.
 XX
 AC ACF68649;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens nucleotide sequence #7116.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 XX
 PN W0200294867-A2.
 XX
 XX 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duclaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 XX WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 2; SEQ ID NO 7116; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification

CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX

SQ Sequence 1191 BP; 346 A; 279 C; 267 G; 299 T; 0 U; 0 Other;

Query Match 58.1%; Score 691.8; DB 10; Length 1191;
 Best Local Similarity 73.8%; Pred. No. 1.7e-209;
 Matches 879; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

Qy 1 ATGTTTGAGAACATTACCGCGCTCTCGCGACCGGATTTGGSCCTGGCGCGATCTGTTT 60
 Db 1 ATGTTTGAGAAAAATCACCGCAGCGCTCGCGACCTTATCTTGGCTTAGCCGATAGTTC 60
 Qy 61 CGTGGCGATGAACGTCCCGGCAAAATTAACCTCGGGATTGGTGCTATAAAGATGAGACG 120
 Db 61 CGTTCGTGATCCTCGTACAAATAAATCAACCTTGGTATCGTGCTATAAAGAGCAACA 120
 Qy 121 GGCAAAACCCCGGTACTGACGCGTGAAAAAGGCTGAACAGTATCTGCTCGAAAAATGAA 180
 Db 121 GGAAAAACCCCGGTCTTGACCCAGTGTAAAAAAGCTGAACAATATTACTTGGAAAAAC 180
 Qy 181 ACCACAAAAAATTACCTCGGCATGACGGCATCCCTGAATTTGGTCTGTCGACTCAGGAA 240
 Db 181 ACAACAAAGAAATTTATCTGCGGCTTAGCGGATTTAGCCGAATTTGCGCGCTAACTCA 240
 Qy 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGTCTCCACGGCAGACT 300
 Db 241 TTAAGTTTGGCAAGATCAACCCAGTTGTACAGATTAACCGCCCGCAGCAAAAGC 300
 Qy 301 CCGGGGGCACTGCGGCACCTACGCGTGGCTGCGGATTTCTTGGCAAAAAATACCAAGCT 360
 Db 301 CCAGCGGTACCGGTGCTTTACGTATTGCTGCGGATTTTCAATGCAACAGACTAATGCT 360
 Qy 361 AAGCGTGTGGGTGAGCAACCCCAAGCTGGCGGAACCATAGACGCTCTTTAACTCTGCA 420
 Db 361 AAACGAGTTTGGATCAGCAACCCCAACCTGGCGCAACCATAAAAACGTTTTTCCGCGCT 420
 Qy 421 GGTCTGGAAGTTCTGTAATACGCTTATTATGATGCGGAAATCACACTCTTGACTTCGAT 480
 Db 421 GGTCTGGAAGTTCTGTAATATAAATACTATGACGCTGAAAAACACGCGCTGAATTTGCA 480
 Qy 481 GCACTGATTAAACAGCCTGAATGAAGCTCAGGCTGGCGACGCTAGTGTGTTCCATGGCTGC 540
 Db 481 GACATGCTGCAAGCCTGTCCGAAGCTCAGGCTGGTGTGTTGTTCTGTTCCACGGCTGC 540
 Qy 541 TGCATAACCCCAACCGGTATCGACCTTACGCTGGAAACAATGGGCAACACATGGGCAACATC 600
 Db 541 TGCACAAATCCGACAGGCATCGATCCAAACCCCGGCAACATGGGCTAAATGGCAGAAATG 600
 Qy 601 TCCGTTGAGAAAGGCTGTTTACCGCTGTTTGAATTCGCTTACCAGGCTTTGCGCGGTG 660
 Db 601 TCTGCGGAGAAAGGCTGTTTGGCTATTTTGAATTTGCTTACCAGGATTTGCCAAGGCG 660
 Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGCTATGCAATAAAGAGCTGATGTT 720
 Db 661 CTAAACGAAGATGAGAGGCGCTACGTATTTTTCGGAATAATCATATGAATGATGTT 720
 Qy 721 GCCAGTTCTTACTCTAAAAAATTGGCCTGTACAAACGAGCGTGTGGCGTTGTACTCTG 780
 Db 721 GCCAGCTCTTACTCCAAAAAATTGGCCTGTACAAATGAACGTTGCGGTGCTGTACTATT 780

QY 781 GTTGTGCGCACAGTGNAAACGGTTGATCGCGCATTCAGCCAAATGAAGCGGCGATTTCG 840
DB 781 GTTGTGCTAGTGACAGTGTATACAGCAGAAAAAGCGTTTACAGCCAAAGTAAAGCGATTATCCGT 840
QY 841 GCTAACTACTCTAAACCCACAGCACAGCGGCGCTTCTGTGTTGCCACCACTCTCTGAGCAAC 900
DB 841 GCTAACTATTCCAAACCCACCGCTCATGTGTGATCTATTGTCTACTACCATTTTGTCAAT 900
QY 901 GATGCGTTACGTCGGATTTTGGGAACAAGAGCTGACTGATATGCGCAGGCGTATTCAGCGT 960
DB 901 GAAGACCTGAAGACGAGCTTGGGAACAGGAAGTACACCACTGCGGCAACGTATCCAGCGT 960
QY 961 ATGCGTCAGTTGTTGCTCAATAGCTGCGAGAAAAAGGCGCAACCGGCACTTCAGCTTT 1020
DB 961 ATGCGTCAACTGTTTGTCAATACCTTGCAGGAAAAAGGCGCAAAACAGGATTTTCAGCTTT 1020
QY 1021 ATCATCAACAGAACGCGATGTTCTCCTTCAGTGGCTGACAAAAGAACAAAGTGTGCGT 1080
DB 1021 ATTATTAGCCAAATGGTATGTTCTCATTCAGTGGCTGACAAAAGAACAAAGTAGAACGT 1080
QY 1081 CTGCGGAAGAGTTTGGCGTATATGCGGTTGCTTCTGTGCGCTAAATGTGCGCGGATG 1140
DB 1081 CTGCGTGATGAGTTTGGTATATACGCTGTGAGTTTCGCTGCTATTAACGTCGCTGCGCTTG 1140
QY 1141 ACACAGATAACATGGTCCGCTGTCGGAAGCGATTTGTGCGACGTGCTGTA 1191
DB 1141 ACGTTAGAAACATGGCTCCACTATGTGAAGCCATTTGTTGACGTACTCTAA 1191

RESULT 10
Continuation (15 of 57) of ACF67367 from base 1400001 (Photorhabdus luminescens nucleoti
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

Fragment Name	Begin	End
ACF67367_01	1	110000
ACF67367_01	100001	210000
ACF67367_02	200001	310000
ACF67367_03	300001	410000
ACF67367_04	400001	510000
ACF67367_05	500001	610000
ACF67367_06	600001	710000
ACF67367_07	700001	810000
ACF67367_08	800001	910000
ACF67367_09	900001	1010000
ACF67367_10	1000001	1110000
ACF67367_11	1100001	1210000
ACF67367_12	1200001	1310000
ACF67367_13	1300001	1410000
ACF67367_14	1400001	1510000
ACF67367_15	1500001	1610000
ACF67367_16	1600001	1710000
ACF67367_17	1700001	1810000
ACF67367_18	1800001	1910000
ACF67367_19	1900001	2010000
ACF67367_20	2000001	2110000
ACF67367_21	2100001	2210000
ACF67367_22	2200001	2310000
ACF67367_23	2300001	2410000
ACF67367_24	2400001	2510000
ACF67367_25	2500001	2610000
ACF67367_26	2600001	2710000
ACF67367_27	2700001	2810000
ACF67367_28	2800001	2910000
ACF67367_29	2900001	3010000
ACF67367_30	3000001	3110000
ACF67367_31	3100001	3210000
ACF67367_32	3200001	3310000
ACF67367_33	3300001	3410000
ACF67367_34	3400001	3510000
ACF67367_35	3500001	3610000
ACF67367_36	3600001	3710000
ACF67367_37	3700001	3810000

WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Query Match 58.1%; Score 691.8; DB 10; Length 110000;
Best Local Similarity 73.8%; Pred. No. 2.1e-208; Indels 0; Gaps 0;
Matches 879; Conservative 0; Mismatches 312;

QY 1 ATGTTTGAGAACATTACCGCCGCTCTGCGGACCCGATTCTGGGCGCTGGCCGATCTGTTT 60
DB 23013 ATGTTTGAGAAATCACCGCAGGCTGCGACCCCTATTCTTGGCTTAGCCGATAGTTT 23072
QY 61 CGTCCGATGAACGTCCTCCGGCAAAATTAACTCTGGGATGGTGTCTATAAAGATGAGCG 120
DB 23073 CGTCTGATCTCGTACAAATAAATCAACCTTGGTATCGGTGTCTATAAAGACGAAACA 23132
QY 121 GCGAAACCCCGTACTACGAGCGTGAAGAGCTGAACAGTACTGCTCCGAAATGAA 180
DB 23133 GGAAGAACCCCGGTTCTGACCAAGTGTAAAAAAGCTGAACAATATTTACTGGAACGAA 23192
QY 181 ACCACCAAAATTTACCTCGGCATTGACGGCATTCCTCGAATTTGGTGCCTGCACCTCAGGAA 240
DB 23193 ACAACAAAGANTTATCTCCGATTAGCGCTTAGCCGANTTTGGCCGCTTAAGGAA 23252
QY 241 CTGCTGTTTGGTAAAGGTAGCGCCCTGATCAATGACAAACGCTGCTCGCACGGCACAGCT 300
DB 23253 TTACTGTTTGGCAAGATCACCCAGTTGTACAGATAAAACGGCCCGCCGACAGCACAAAGC 23312
QY 301 CGGGGGGACCTGGCGCATACGCTGGCTGCGGATTTCTTGGCAAAATATACCGGTT 360
DB 23313 CCAGCGGTACCGGTTTACGTTATGCTGGGATTTTCATTCGCAACAGACTAATGCT 23372
QY 361 AAGCGTGTGTGGTGAGCAACCCAAAGCTGGCCGAAACCAATAAGAGCGTCTTTAACTCTGCA 420
DB 23373 AAACGAGTTTGGATCAGCAACCCCAACCTGGCCCAACCATATAAAGCGTTTTTCCGCGCT 23432
QY 421 GGTCTGGAAAGTTCTGTAATACGCTTATTATGATGGGAAAAATCACACTCTTGACTTCGAT 480
DB 23433 GGTCTGGAAAGTCTGTGAATATAAATACTATGACGCTGAAAAACACGCGCTGAATTTTCGAA 23492
QY 481 GCACTGATTAACAGCCTGAATGAAGCTCAGGCTGGCGCAGTAGTGCTGTTTCCATGCTGC 540
DB 23493 GACATGCTGGCAAGCGCTGCCAAGCTCAGGCTGGTGATGTTGTTCTGTTTCCACGCGTGC 23552
QY 541 TGCATAACCCCAACCGGTATCGACCTACGCTGGAAACAATGGCAACACACTGGCACAACTC 600
DB 23553 TGCACAAATCCGACAGGCATCAATCCAAACCCGGCACATGGGCTAACTGGCAGAAATG 23612
QY 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTTGAATTCGCTTACAGGGTTTGGCCCGTGGT 660
DB 23613 TCTGCGGAGAAAGGCTGGTTGCTTATTTTGTATTCGCTTACCAAGGATTTGCCAAGGGC 23672
QY 661 CTGGAAGAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCAATAAAGAGCTGATGTT 720
DB 23673 CTAACGAAAGATGACAGAGGCGCTACGTTATTTTTCGAAAAATCATATAAGACTGATGTT 23732
QY 721 GCCAGTTCTACTCTATAAAACTTTTGGCCTGTACAAACGAGCGTGTGGCGCTTGTACTCTG 780

Db 23733 GCCAGCTCTTACTCCAAAATTTGGCGCTGTACAAATGAAGTGTGCGTCTGTACTATT 23792
 Qy 781 GTTGCTGCCGACAGTGAACCGTTGATCGCGCATTTAGCCAAATGAAGCGGCGATTCGC 840
 Db 23793 GTTGCTAGTACAGTGATACGACGAGAAAAGCGTTACGCCAAGCTAAGCGATATCCGT 23852
 Qy 841 GCTAACTACTTAAACCCAGCAGACACGCGGCTTCTGTTGTCACCATCTCTGAGCAAC 900
 Db 23853 GCTAACTATTCCAAACCCAGCGCTCATGTCATCTATTGTCACCTACCAATTTGTCAAT 23912
 Qy 901 GATGCGTTAGCTGCGATTGGACACAGAGCTGACTGATATGCGCCAGCGTATTCAGCGT 960
 Db 23913 GAAGACCTGAAGCAGCTTGGGAACAGAACTGACCACTGCGCGAAGCTATCCAGCGT 23972
 Qy 961 ATGCGTCAGTTGTTGTCATACGCTGCAGAAAAGCGCAACCGCGACTTTCAGCTTT 1020
 Db 23973 ATGCGTCAGTTGTTGTCATACCTGACGAAAAGCGCAACCGCGACTTTCAGCTTT 24032
 Qy 1021 ATCATCAACAGAACCGCATGTTCTCTCAGTGGCTGACAAAAGAACAAAGTCTGCGT 1080
 Db 24033 ATTATTAGCCAAAATGATGTTCTCTATTGAGTGGCTGACAAAAGAACAAAGTAGAACGT 24092
 Qy 1081 CTGCGGAGAGCTTGGCGTATATGCGTGTCTTCTGCGTAAATGTGGCGGGATG 1140
 Db 24093 CTGCGTGATGAGTTGGTATATAGCTGTGAGTTCCGCTGCTATTAACGTCGCTGGCTG 24152
 Qy 1141 ACACCATCAATGCTGCGCTGTGCGAAGCGATTGTGCGAGTGTGTTAA 1191
 Db 24153 ACGTTAGAAAACATGGCTCCACTATGTGAAGCCATTGTTGAGTACTCTAA 24203

RESULT 11

ACF65381
 ID ACF65381 standard; DNA; 249878 BP.
 XX ACF65381;
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens nucleotide sequence #34.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 XX
 PN WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 XX WPI; 2003-148459/14.
 DR
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 1; SEQ ID NO 34; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens

CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 249878 BP; 75946 A; 56652 C; 49249 G; 68029 T; 0 U; 2 Other;
 Query Match 58.1%; Score 691.8; DB 10; Length 249878;
 Best Local Similarity 73.8%; Pred. No. 3.3e-208;
 Matches 879; Conservative 0; Mismatches 312; Indels 0; Gaps 0;
 Qy 1 ATGTTTGGAGAACATTACCGCGCTCTCTGCGACCCGATTTCTGGGCGCTGGCGGATCTCTTT 60
 Db 1334 ATGTTTGGAGAAATACCGCGCGCTCTGCGACCCCTATTCTTGGCTAGCCGATAGTTTC 1393
 Qy 61 CGTGCCGATGAACGTCCTCGGCAAAATTAACCTCGGATTTGCTATATAAGATGAGACG 120
 Db 1394 CGTTCGTATCTCGTACAAATAAATCAACCTTGGTATCGGTCTCTATAAAGACGAAACA 1453
 Qy 121 GGCAAAACCCGCTACTGACCGCTGAAAAGGCTGACAGTATCTGCTCGGAAATGAA 180
 Db 1454 GGAAAAACCCCGGTTCTGACCAAGTGTAAAAAAGCTGAAATATTCTTGGAAAAACGAA 1513
 Qy 181 ACCACCAAAATTAACCTCGGCAATTCGCGCATCCCTGAAATTTGCTGCTGCATCAGGAA 240
 Db 1514 ACAACAAGATTTATCTGCCGATTAGCGGCTTAGCCGAATTTGCGCGGTAACCTCAAGAA 1573
 Qy 241 CTGCTCTTTGTTAAAGTAGCGCCCTGATCAATGACAAAGCTGCTCGCAGGACAGACT 300
 Db 1574 TTACTGTTTGGCAAGATCACCCAGTTGTACAGATAAACCGCCGCGCAGACACAAAGC 1633
 Qy 301 CCGGGGGCACTGGCGCACTAGCGTGGCTGCGGATTTCTTGGCAAAAATATACGAGTT 360
 Db 1634 CCAGGCGGTACCGGTGCTTTACGATTTGCTGCGGATTTTCAATGCCAAACAGACTAATGCT 1693
 Qy 361 AAGCGTGTGTGGGTGAGCAACCCCAAGCTGCGCCGAAACATAAGAGCGTCTTTAACTCTGCA 420
 Db 1694 AACAGATTTGGATCAGCAACCCACCTGGCCAAACATATAAAACGTTTTTTCGCGCGCT 1753
 Qy 421 GGTCTGGAAGTTGCTGAATACCGCTTATATGATGCGGAAATACACTCTTTGATCTCGAT 480
 Db 1754 GGTCTGGAAGTCTGTGAATATAAATACTATGACGCTGAAAAACACGCGCTGAATTTTCGAA 1813
 Qy 481 GCACTGATTAACAGCCTGATGAGCTCAGGCTGGCGCAGCTAGTGTGTTTCCATGGCTGC 540
 Db 1814 GACATGCTGGCAAGCCTGTCGGAAGCTCAGGCTGGGTGATGTTGTTCTGTTCCACGCGTGC 1873
 Qy 541 TGCATATAACCAACCGGTATCGACCTTACGCTGGAACAAATGGCAAAACACTGGCACAACCTC 600
 Db 1874 TGCCACATCCGACAGGCTATGATCCAAACCCGCGCAATGGGCTAACTGGCAGAAATG 1933
 Qy 601 TCCGTTGAGAAAGGCTGGTTACCGCTGTTGACTTCCGTTACAGGGTTTTCGCGGTGGT 660
 Db 1934 TCTCGGAGAAAGGCTGGTTGCTATTTTGTATTTCCGTTACCAAGGATTTGCAAGGCG 1993
 Qy 661 CTGGAAGAAGATGCTGAAGGACTGCGCGCTTTTCGCGGCTATGCTATAAGAGCTGATGTT 720
 Db 1994 CTAACGAAGATGACAGAGGCGCTACGATATTTTTCGGAATAATCATAATGAACGTATGTT 2053

Db 541 TGCATTAACCCAAACCGGTATCGATCCAAACGGCTGAACAGTGGCAGAACTGGCAGATTTA 600
Qy 601 TCCGTTGAGAAAGCTGTTACCGCTGTTTACCTTTCGCTTACCAAGGTTTTCGCCGCGGT 660
Db 601 TCTGCGCAATGCTGGCTGCTGTTGATTTTCGCTTATCAGGTTTTCGCAAGT 660
Qy 661 CTGGAAGAAGATGCTGAAGGATCGCGCTTTTCGCGCTATGCAATAAGAGCTGATTGTT 720
Db 661 TTAGATGAAGATGAGAGGCTCGTATTTTACTAAAAACCAATAATGAATGATCGTT 720
Qy 721 GCCAGTCTTACTTAATAACTTTGGCTGTGACAGGCGTGTGGCGCTGTACTCTG 780
Db 721 GCCAGCTCTTACTCAAAAAATTTTGGCTGTGACAAATGAGCGGTGTAGTGCTGCACCATC 780
Qy 781 GTTGCTGCCAGACAGTGAACCGTTGATCGCGCATTTACGCCAAATGAAGCGGCGATTTCGC 840
Db 781 GTTGCACTGACAGCATACCGCAGAAAGAGCATTCAGTCAGGCGAAATCTATTGTCGT 840
Qy 841 GCTAACTACTTAACCCACGACACCGCGCTTCTGTTGTCACCATCTCTGAGCAAC 900
Db 841 ACTAACTACTTAACCCACGCGCACCGCGCTTCTGTTGTTACTACAATTTCTGCCAAT 900
Qy 901 GATCGCTTACGTGCGATTTGGGAACAAGAGTGTACTATATGCCAGCGATTTCACGCT 960
Db 901 GACGAATTTAAAGCAGAAATGGATTCAAGAACTGGCGACTATGCGTGAACGCAATTCGTCGT 960
Qy 961 ATGCGTCAGTGTTCGTCATACGCTGCAGGAAAAGCGCAACCGCGACTTCAGCTTT 1020
Db 961 ATGCGTCAGTGTTCGTCATACGCTGCAGGAAAAGCGCAACCGCGACTTCAGCTTT 1020
Qy 1021 ATCATCAACAGACGCGCATGTTCTCTCAGTGGCCCTGCAAAAGAACCAAGTGTCTCGT 1080
Db 1021 ATTATTAGCAAAATGATGTTTTCATTCAGTGTCTGACTAAGAAACAAGTTGATCGT 1080
Qy 1081 CTGCGCAAGATTTGGCGTATATGCGGTGCTTCTGCTGCGTAAATGTGGCGGATG 1140
Db 1081 CTGCGTGAAGAAATTCGCTATTTATGCGGTGAGCTCTGGCGGCAATTAACGTAGCGGCTTA 1140
Qy 1141 ACACCAATCAATGCTGCTGCGTGTGCAAGCGGATTCGTGCGAGTCT 1187
Db 1141 ACACCTGAGAACATGGTTCTCTGTGGAAGCGATTTGTCAGTACT 1187

RESULT 13
ADS45488
ID ADS45488 standard; cDNA; 966 BP.
AC ADS45488;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #231.
XX

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX

OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOV/) CAO Y.
PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 23918; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 966 BP; 258 A; 221 C; 236 G; 251 T; 0 U; 0 Other;

Query Match 42.8%; Score 510; DB 13; Length 966;
Best Local Similarity 70.5%; Pred. No. 1.7e-151;
Matches 681; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

Qy 223 GGTGCTGCTACTCAGGAACCTGCTTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGT 282
Db 1 GCGCCGGGTGATTGAGGAACCTGCTTTTGGTAAAGTAGCGCCCTGATCAATGACAAACGT 60
Qy 283 GCTCGCACGGCACAGACTCCCGGGGGCACCTGGCGGCACCTAGCGGTGGCTGCGATTTCCTG 342
Db 61 GCGCGTACCGTACAAAGCCAGCGGNAACAGGCGCACTGCTGCTGCGGTGATTATT 120
Qy 343 GCAAAAAATACCAAGCGTTAAGCGTGTGTGGTGAGCAACCAAGCTGGCGCAACATAAG 402
Db 121 GCCAAGCAAACTAATGCCAAACGCTGTTGGATCAGTAACCAACCTGGCCAAACATAAA 180
Qy 403 AGCGTCTTTAACTCTCGAGTCTGGAAGTTCGTGAATACGCTTATTATGATCGGAAAT 462
Db 181 GGTGTTTTTCCAGCGCAGGTTTAGAGATCCCGGAATATAACTATTACGATGCGAGAAAA 240
Qy 463 CACACTCTTGACTTCGATGCACCTGATTAAACAGCTGAATCAAGCTCAGGCTGGCGAGCTA 522
Db 241 CACGCTCTGAATTTGAGAGCATGCTGGCAAGCGCTGTCTGAGCACAAGCGGGTGATTT 300
Qy 523 GTGCTGTTCATGGCTGTGCCATAACCCCAACCGGTATCGACCTACGCTGGAAACAATGG 582
Db 301 GTTCTGTGACGCTTGTGTCGCATAACCCGACCGGTATCGATCTCTACTCCAGAACAGTGG 360
Qy 583 CAAACACTGGCAACAATCTCTCCGTTGAGAAAGGCTGTTACCGCTGTTGACTTCGCTTAC 642
Db 361 CAAAAAATGGCAGATTTATCTCGCGCAATGGCTGGCTGCTGTATTTCGATTTTCGCTTAT 420

ID AAS90082 standard; cDNA; 3222 BP.
 AC AAS90082;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #25886.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG25895.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 25886; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3222 BP; 727 A; 880 C; 863 G; 752 T; 0 U; 0 Other;
 Query Match 41.7%; Score 496.2; DB 5; Length 3222;
 Best Local Similarity 99.4%; Pred. No. 8.2e-147;
 Matches 498; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 687 CGCTTTCCGGCTATGCATAAAGAGCTGATTTGGCAGTTCTTCTTAAACCTTTGG 746
 DB 1425 CCCTTTCCGGCTATGCATAAAGAGCTGATTTGGCAGTTCTTCTTAAACCTTTGG 1484
 QY 747 CCTGTACACGACGCTTGGCGCTTGTACTCTGTTGCTGCCAGAGTGAACCGTTGA 806
 DB 1485 CCTGTACACGACGCTTGGCGCTTGTACTCTGTTGCTGCCAGAGTGAACCGTTGA 1544
 QY 807 TCGCGCATTCAGCCAAATGAAGCGGCGATTTCGCGTAACTACTCTAAACCCAGCACA 866

Db	1545	TCGCGCATTCAGCCAAATGAAGCGGCGATTTCGCGTAACTACTCTAAACCCAGCACA	1604
QY	867	CGGCGCTTCTGTTGTTGGCCACCATCTGAGCAACGATGCGTTACGTGCGATTTGGGAACA	926
Db	1605	CGGCGCTTCTGTTGTTGGCCACCATCTGAGCAACGATGCGTTACGTGCGATTTGGGAACA	1664
QY	927	AGAGCTGACTGATATGCGCCAGCGTATTTCAGCGTATGCGTCACTTGTTCGTCATACGCT	986
Db	1665	AGAGCTGACTGATATGCGCCAGCGTATTTCAGCGTATGCGTCACTTGTTCGTCATACGCT	1724
QY	987	GCAGGAAAAAGCGGCAAAACCGGACCTTCAGCTTTATCATCAACAGACGGCATGTTCTC	1046
Db	1725	GCATGAAAAAGCGGCAAAACCGGACCTTCAGCTTTATCATCAACAGACGGCATGTTCTC	1784
QY	1047	CTTCAGTGGCCTGACAAAAAGAACCAAGTGTGCGTCTGCGGAAAGAGTTTGGCGTATATGC	1106
Db	1785	CTTCAGTGGCCTGACAAAAAGAACCAAGTGTGCGTCTGCGGAAAGAGTTTGGCGTATATGC	1844
QY	1107	GGTTGCTTCTGTTGCGTAAATGTGCGCGGATGACACCAAGATGCTCCGCTGTG	1166
Db	1845	GGTTGCTTCTGTTGCGTAAATGTGCGCGGATGACACCAAGATGCTCCGCTGTG	1904
QY	1167	CGAAGCGATTGTGGCAGTGCT	1187
Db	1905	CGAAGCGATTGTGGCAGTGCT	1925

Search completed: March 15, 2006, 03:02:48
 Job time : 1397 secs

This Page Blank (uspto)

CC lysine produced can be used as a fodder additive

XX Sequence 1331 BP; 330 A; 340 C; 350 G; 311 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 1191; DB 2; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAAGAACATTAACCGCGCTCTCGCGACCGCGATTCCTGGCCCTGGCCGATCTGTTT 60
DB 10 ATGTTTGAAGAACATTAACCGCGCTCTCGCGACCGCGATTCCTGGCCCTGGCCGATCTGTTT 69

QY 61 CGTGCCGATGAACGTCCTCCCGCAAAATTAACCTCGGATTCGCTCTATAAGATGAGACG 120
DB 70 CGTGCCGATGAACGTCCTCCCGCAAAATTAACCTCGGATTCGCTCTATAAGATGAGACG 129

QY 121 GSCAAAACCCCGTACTACGACGCTGGAAGGCTGAACAGTATCTGCTCGAAATGAA 180
DB 130 GSCAAAACCCCGTACTACGACGCTGGAAGGCTGAACAGTATCTGCTCGAAATGAA 189

QY 181 ACCACAAAATTAACCTCGGATTCGCGCATCCCTGAATTTGGTGGCTGCACTCAGGAA 240
DB 190 ACCACAAAATTAACCTCGGATTCGCGCATCCCTGAATTTGGTGGCTGCACTCAGGAA 249

QY 241 CTGCTGTTTGTAAAGTAGCGCTCTGATCAATGACAAACGCTGCTCGCACGCACT 300
DB 250 CTGCTGTTTGTAAAGTAGCGCTCTGATCAATGACAAACGCTGCTCGCACGCACT 309

QY 301 CCGGGGGCACTGGGCACTACGCGTCTCGGATTTCTGGCAAAATTAACAGCGTT 360
DB 310 CCGGGGGCACTGGGCACTACGCGTCTCGGATTTCTGGCAAAATTAACAGCGTT 369

QY 361 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCAATAGAGCGCTTTTAACTCTGCA 420
DB 370 AAGCGTGTGGGTGAGCAACCCAGCTGGCGCAACCAATAGAGCGCTTTTAACTCTGCA 429

QY 421 GGTCTGAAGTTGCTGAATACGCTTATTATGATCGGAAATACACCTCTGACTTCGAT 480
DB 430 GGTCTGAAGTTGCTGAATACGCTTATTATGATCGGAAATACACCTCTGACTTCGAT 489

QY 481 GCACGTATTAACGCTGAATGAGCTCAGGCTGGCGACGTAGTGTCTTCCATGGCTGC 540
DB 490 GCACGTATTAACGCTGAATGAGCTCAGGCTGGCGACGTAGTGTCTTCCATGGCTGC 549

QY 541 TGCATAACCCACCGGTATCGACCTACGCTGGAACAAATGGCAACACTGGCACTC 600
DB 550 TGCATAACCCACCGGTATCGACCTACGCTGGAACAAATGGCAACACTGGCACTC 609

QY 601 TCGGTGAGAAAGGCTGGTTACCGCTGTTTACCTGCTTACAGGGTTTGGCCGTGT 660
DB 610 TCGGTGAGAAAGGCTGGTTACCGCTGTTTACCTGCTTACAGGGTTTGGCCGTGT 669

QY 661 CTGGAAGAAGATGCTGAAGGATGCGCGCTTTTGGCGCTATGATGAAGAGCTGATTCGT 720
DB 670 CTGGAAGAAGATGCTGAAGGATGCGCGCTTTTGGCGCTATGATGAAGAGCTGATTCGT 729

QY 721 GCCAGTCTCTACTTAAACCTTTGGCGCTGTAACAACGAGCGTTTGGCGCTTGTACTCTG 780
DB 730 GCCAGTCTCTACTTAAACCTTTGGCGCTGTAACAACGAGCGTTTGGCGCTTGTACTCTG 789

QY 781 GTTGTGCGCAGTGAACACCGTTGATCGCGCATTCAGCCCAATGAAGCGCGGATTCGC 840
DB 790 GTTGTGCGCAGTGAACACCGTTGATCGCGCATTCAGCCCAATGAAGCGCGGATTCGC 849

QY 841 GCTAACTACTTAACCCACAGCAACGCGCTCTCTGTTGTCACCATCTCAGCAAC 900
DB 850 GCTAACTACTTAACCCACAGCAACGCGCTCTCTGTTGTCACCATCTCAGCAAC 909

QY 901 GATCGCTTACGTCGATTTGGGAACAAGCTGATGATATGCGCCAGCGTATTCAGCGT 960
DB 910 GATCGCTTACGTCGATTTGGGAACAAGCTGATGATATGCGCCAGCGTATTCAGCGT 969

QY 961 ATGCGTCAGTTGTTGCTCAATACGCTGCAGAAAGCGCAACCCGCACTTCAGCTTT 1020

DB 970 ATCGTCAGTTGTTGTCATACGCTGACGAAAAAGCGCAACCCGACCTTCAGCTTT 1029
QY 1021 ATCATCAACAGAACGCGCATGTTCTCTTCAAGTGGCCTGACAAAGAACAGTCTCGGT 1080
DB 1030 ATCATCAACAGAACGCGCATGTTCTCTTCAAGTGGCCTGACAAAGAACAGTCTCGGT 1089
QY 1081 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTCGCGGATG 1140
DB 1090 CTGCGCGAAGAGTTTGGCGTATATGCGGTGCTTCTGTCGCGTAAATGTCGCGGATG 1149
QY 1141 ACACAGATAAATCAATGCTCGCTGCGAAGCGATTTGTCGAGTGTCTAA 1191
DB 1150 ACACAGATAAATCAATGCTCGCTGCGAAGCGATTTGTCGAGTGTCTAA 1200

RESULT 5

AN71109
ID AN71109 standard; DNA; 3659 BP.
XX
AC AN71109;
XX
DT 01-JAN-1980 (first entry)
XX
DE pheA aroF aspC operon in plasmid pME219.
XX
KW alpha-amylase; feedback inhibition; amino acid synthesis;
XX
KW composite plasmid; ss.
XX
OS Bacillus licheniformis.

Key Location/Qualifiers
CDS 69..1241
/*tag= a
CDS 1259..2329
/*tag= b
CDS 2344..3534
/*tag= c

W08700202-A.

15-JAN-1987.

24-JUN-1986; 86WO-US001353.

24-JUN-1985; 85US-00747732.

(NUTR-) NUTRASWEET CO.

Edwards MR, Taylor PP, Hunter MG, Fotheringh IG;

WPI; 1987-021998/03.

P-PSDB; AAP70752, AAP71677, AAP71678.

Composite plasmids contg. multiple genes in transcriptional units -
useful for prodn. of aminoacid(s), esp. L-phenylalanine and L-tyrosine.
Disclosure; Page 38; 57pp; English.

This sequence may be inserted into a composite plasmid and used for the
production of amino acids. See also AN71053-55, AN71107, AN71109,
AN71111 and AAP70696-97 and AAP70750, AAP70752 and AAP70754

Sequence 3659 BP; 936 A; 911 C; 947 G; 865 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 1191; DB 1; Length 3659;

Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTGAAGAACATTAACCGCGCTCTCGCGACCGCGATTCGCGCTGGCGGATCTGTTT 1020
DB 2344 ATGTTTGAAGAACATTAACCGCGCTCTCGCGACCGCGATTCGCGCTGGCGGATCTGTTT 1029

QY 61 CGTCCGATGAACGCTCCGGCAAAATTAACTCCGGATTGGTGTCTATAAAGATGAGACG 120
Db 2404 CGTCCGATGAACGCTCCGGCAAAATTAACTCCGGATTGGTGTCTATAAAGATGAGACG 2463
QY 121 GCGAAACCCGGTACTGACAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 180
Db 2464 GCGAAACCCGGTACTGACAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 2523
QY 181 ACCACCAAAATTAACCTCGGCATTGACCGCATCCCTGAAATTTGGTTCGCTGCACTCAGGAA 240
Db 2524 ACCACCAAAATTAACCTCGGCATTGACCGCATCCCTGAAATTTGGTTCGCTGCACTCAGGAA 2583
QY 241 CTGCTGTTGGTAAAGGTAGCCCTGTGATCAATGACAAACGTCGTCGACGCAAGACT 300
Db 2584 CTGCTGTTGGTAAAGGTAGCCCTGTGATCAATGACAAACGTCGTCGACGCAAGACT 2643
QY 301 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGGATTTCTTGGCAAAATACACGAGCTT 360
Db 2644 CCGGGGGGCACTGGCGCACTACGCTGGCTGCGGATTTCTTGGCAAAATACACGAGCTT 2703
QY 361 AAGCGTGTGTGGTGAACCAACCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 420
Db 2704 AAGCGTGTGTGGTGAACCAACCAAGCTGGCGCAACCAATAGAGCGTCTTTAACTCTGCA 2763
QY 421 GGTCTGGAAGTTGCTGAATACGCTTATTATGATGCGGAAATACACTCTTGCATTCGAT 480
Db 2764 GGTCTGGAAGTTGCTGAATACGCTTATTATGATGCGGAAATACACTCTTGCATTCGAT 2823
QY 481 GCACTGATTAAAGCCTGAATGAGCTCAGGCTGGCGAGTGTGTTTCATCGCTGC 540
Db 2824 GCACTGATTAAAGCCTGAATGAGCTCAGGCTGGCGAGTGTGTTTCATCGCTGC 2883
QY 541 TGCATAACCCCAACCGGTATCGACCTACGCTGGAACAATGCAACACTGCGCAACACTC 600
Db 2884 TGCATAACCCCAACCGGTATCGACCTACGCTGGAACAATGCAACACTGCGCAACACTC 2943
QY 601 TCCGTTGGAAGGCTGTTACCGCTGTTGACTTCGCTTACCAGGGTTTTGCCCGGTGT 660
Db 2944 TCCGTTGGAAGGCTGTTACCGCTGTTGACTTCGCTTACCAGGGTTTTGCCCGGTGT 3003
QY 661 CTGGAAGAAGATGCTGAAGCACTGCGGCTTTCGGGCTATGCAATAAGACTGATTGTT 720
Db 3004 CTGGAAGAAGATGCTGAAGCACTGCGGCTTTCGGGCTATGCAATAAGACTGATTGTT 3063
QY 721 GCGAGTCTCTACTTAAAACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 780
Db 3064 GCGAGTCTCTACTTAAAACTTTGGCTGTACACGAGCGTGTGGCGCTTGTACTCTG 3123
QY 781 GTTGTGCGCAGTGAACCCGTTGATCGCGATTGACGCAATGAAAGCGCGATTGCG 840
Db 3124 GTTGTGCGCAGTGAACCCGTTGATCGCGATTGACGCAATGAAAGCGCGATTGCG 3183
QY 841 GCTAACTACTTAAACCCACGACGACGCGGCTTCTGTTGTTGCCACCATCTGAGCAAC 900
Db 3184 GCTAACTACTTAAACCCACGACGACGCGGCTTCTGTTGTTGCCACCATCTGAGCAAC 3243
QY 901 GATGCGTTACGTGCGATTGGGAAACAGAGCTGACTGATGATGCGCAGCGTATTGAGCGT 960
Db 3244 GATGCGTTACGTGCGATTGGGAAACAGAGCTGACTGATGATGCGCAGCGTATTGAGCGT 3303
QY 961 ATGCGTCACTGTTGCTCAATACGCTGACAGGAAAGGCGCAACCGGACCTTCAGCTTT 1020
Db 3304 ATGCGTCACTGTTGCTCAATACGCTGACAGGAAAGGCGCAACCGGACCTTCAGCTTT 3363
QY 1021 ATCATCAACAGAACGGCAATGTTCTCTTCACTGAGCTGACAAAGAACAGTGTGCGT 1080
Db 3364 ATCATCAACAGAACGGCAATGTTCTCTTCACTGAGCTGACAAAGAACAGTGTGCGT 3423
QY 1081 CTGCGCGAAGAGTTGGCGGTATATGCGGTTGCTTCTGTCGCGTAAATGTGCCCGGATG 1140
Db 3424 CTGCGCGAAGAGTTGGCGGTATATGCGGTTGCTTCTGTCGCGTAAATGTGCCCGGATG 3483
QY 1141 ACACGAGATAACATGGCTCCGCTGTCGGAAGCGATTGTGGCAGTGTCTGTAA 1191

File ,

Db 3484 ACACGAGATAACATGGCTCCGCTGTGCGAAGCGATTGTGCGAGTGTCTGTAA 3534
RESULT 6
AAS46273/C
ID AAS46273 standard; DNA; 14759 BP.
XX AAS46273;
XX 18-DEC-2001 (first entry)
XX DNA encoding novel mar regulated protein (NIMR) #42.
XX mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.
XX Escherichia coli.
XX WO200170776-A2.
XX 27-SEP-2001.
XX 08-MAR-2001; 2001WO-US007478.
XX 10-MAR-2000; 2000US-0188362P.
XX (TUFT) TUFTS COLLEGE.
XX Levy SB, Barbosa TM, Alekshun MN;
XX WPI; 2001-602769/68.
XX P-PSDB; AAU29374.
XX Identifying compounds that modulate a newly identified mar regulated
XX polypeptide activity, useful as antimicrobial compounds, involves
XX contacting the polypeptide with a test compound.
XX Disclosure; Page 477-485; 526pp; English.
XX The invention relates to a method of identifying compounds that modulate
XX a newly identified mar regulated (NIMR) polypeptide activity. The method
XX comprises contacting an NIMR polypeptide with a test compound under
XX interaction conditions, determining the ability of the compound to
XX modulate the activity or expression of the polypeptide, and selecting the
XX modulators. NIMR nucleic acids and polypeptides are used in the treatment
XX of microbial infections, and in screening for modulators of NIMR
XX expression and activity. These modulators can be used to reduce the
XX infectivity of a microbe on a surface, and the virulence of a microbe in
XX a subject suffering from an infection. AAS46232-AAS46278 represent
XX Escherichia coli NIMR coding sequences of the invention
XX SQ Sequence 14759 BP; 3703 A; 3840 C; 3713 G; 3503 T; 0 U; 0 Other;
Query Match 100.0%; Score 1191; DB 4; Length 14759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTGAAGACATTACCGCGCTCTCCGACCCGATTTCTGGGCTTGGCCGATCTGTTT 60
Db 2755 ATGTTTGAAGACATTACCGCGCTCTCCGACCCGATTTCTGGGCTTGGCCGATCTGTTT 2696
QY 61 CGTCCGATGAACGCTCCCGCAAAATTAACCTCCGGATTGGTGTCTATAAAGATGAGACG 120
Db 2695 CGTCCGATGAACGCTCCCGCAAAATTAACCTCCGGATTGGTGTCTATAAAGATGAGACG 2636
QY 121 GCGAAACCCGGTACTGACAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 180
Db 2635 GCGAAACCCGGTACTGACAGCTGAAAGGCTGAACAGTATCTGCTCGAAATGAA 2576
QY 181 ACCACCAAAATTAACCTCGGCATTGACCGCATCCCTGAAATTTGGTTCGCTGCACTCAGGAA 240
Db 2575 ACCACCAAAATTAACCTCGGCATTGACCGCATCCCTGAAATTTGGTTCGCTGCACTCAGGAA 2516

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	2045	100.0	396	1	AAP17678
2	2045	100.0	396	2	AAW69553
3	2045	100.0	396	7	AAD833285
4	2045	100.0	396	8	ADN18132
5	2045	100.0	396	9	ADW95404
6	1931	94.4	402	7	ABO65556
7	1682	82.2	397	6	ABM69171
8	1654	80.9	396	8	AD342695
9	1318	64.4	322	8	ADN17578
10	1317	64.4	423	6	ABP80003
11	1311	64.1	397	5	AAU73004
12	1310	64.1	397	8	ADF08222
13	1005.5	49.2	394	8	AD322791
14	996	48.7	396	8	ADN24844
15	996	48.7	398	8	ADN222085
16	982	48.0	397	8	ADN25771
17	982	48.0	413	7	ABO80353
18	975	47.7	397	8	AD326795
19	975	47.7	397	8	AD327175
20	975	47.7	398	8	AD326427
21	947	46.3	398	8	AD324992
22	924	45.2	395	8	ADN26805
23	917	44.8	395	8	ADN26569
24	917	44.8	420	8	ADY07798

Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKPTVLTSVKKAQYLLENE 60
 Qy 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTGGTGALRVAADFLAKNTSV 120
 Db 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTGGTGALRVAADFLAKNTSV 120
 Qy 121 KRWVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
 Db 121 KRWVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180
 Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLDFDFAQGFARGLEDEDAEGLRAFAAMHKELIV 240
 Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLDFDFAQGFARGLEDEDAEGLRAFAAMHKELIV 240
 Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
 Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300
 Qy 301 DALRAIWEQELTDMRQRIQRMROLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360
 Db 301 DALRAIWEQELTDMRQRIQRMROLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360
 Qy 361 LREDFGVYAVASGRVNVAGTDPDNMAPLCEAIVAVL 396
 Db 361 LREDFGVYAVASGRVNVAGTDPDNMAPLCEAIVAVL 396

RESULT 2

AAW69553
 ID AAW69553 standard; protein; 396 AA.

AC AAW69553;

DT 13-OCT-1998 (first entry)

DE Escherichia coli aspC protein.

XX Brevibacterium lactofermentum; lysC; L-lysine; coryneform bacterium;
 KW aspartokinase; feedback inhibition; dihydrodipicolinate reductase;
 KW diaminopimelate decarboxylase; aspartate aminotransferase.

XX Escherichia coli.

XX EP854189-A2.

XX 22-JUL-1998.

XX 05-DEC-1997; 97EP-00121443.

XX 05-DEC-1996; 96JP-00325659.

XX (AJIN) AJINOMOTO CO INC.

XX Araki M, Sugimoto M, Yoshihara Y, Nakamatsu T;

XX WPI; 1998-379060/33.

DR N-PSDB; AAV40259.

XX Recombinant DNA autonomously replicable in coryneform bacteria - used to
 PT produce L-lysine, codes for e.g. aspartokinase, dihydrodipicolinate
 PT reductase and synthase and di-amino-pimelate decarboxylase.

XX Claim 6; Page 38-39; 59pp; English.

XX The present invention describes a recombinant DNA autonomously replicable
 CC in cells of coryneform bacteria (CB), comprising a DNA sequence coding
 CC for an aspartokinase (AK) in which feedback inhibition by L-lysine and L-
 CC threonine is desensitized, a DNA sequence coding for a
 CC dihydrodipicolinate reductase (DHP), a DNA sequence coding for
 CC dihydrodipicolinate synthase (DHPs), a DNA sequence coding for
 CC diaminopimelate decarboxylase (DAMD) and a DNA sequence coding for
 CC aspartate aminotransferase (AAT). The present sequence represents aspC
 CC from Escherichia coli. The DNA and related products from the present

CC invention, can be used for improving L-lysine productivity by CB. The L-
 CC lysine produced can be used as a fodder additive

SQ Sequence 396 AA;

Query Match 100.0%; Score 2045; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 2e-197;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKPTVLTSVKKAQYLLENE 60

Db 1 MFENITAAPADPILGLADLFRADERPGKINLGIVYKDETKPTVLTSVKKAQYLLENE 60

Qy 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTGGTGALRVAADFLAKNTSV 120

Db 61 TTKNYLGIDGIPFGRCTOELLFGKGSALINDKRAATAQTGGTGALRVAADFLAKNTSV 120

Qy 121 KRWVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180

Db 121 KRWVSNPSPNKHVSFNSAGLEVREYAYDAENHTLDFDALINSLNEAAGDVVLPFGC 180

Qy 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLDFDFAQGFARGLEDEDAEGLRAFAAMHKELIV 240

Db 181 CHNPTGIDPTLEQWOTLAQLSVEKGLPLDFDFAQGFARGLEDEDAEGLRAFAAMHKELIV 240

Qy 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

Db 241 ASSYSKNFGLYNERVGACTLVAADSETVDRAFSQMKAAIRANYSNPPAHGASVVATILSN 300

Qy 301 DALRAIWEQELTDMRQRIQRMROLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360

Db 301 DALRAIWEQELTDMRQRIQRMROLFVNTLQEKGNRDFSFIIKONGMFSFGLTKQVLR 360

Qy 361 LREDFGVYAVASGRVNVAGTDPDNMAPLCEAIVAVL 396

Db 361 LREDFGVYAVASGRVNVAGTDPDNMAPLCEAIVAVL 396

RESULT 3

AD883285

ID AD883285 standard; protein; 396 AA.

XX AC AD883285;

XX DT 04-DEC-2003 (first entry)

XX DE Escherichia coli aspartate aminotransferase.

XX aspartate aminotransferase; threonine; fermentation; enzyme.

XX Escherichia coli.

XX WO2003072786-A1.

XX 04-SEP-2003.

XX 25-FEB-2003; 2003WO-JP002067.

XX 27-FEB-2002; 2002RU-00104983.

XX (AJIN) AJINOMOTO CO INC.

XX Akhverdian VZ, Savrasova EA, Kaplan AM, Lobanov AO, Kozlov YI;

XX WPI; 2003-721777/68.

DR N-PSDB; ADB83284.

XX Industrial production of L-threonine by fermentation using Escherichia
 PT modified to enhance aspartate aminotransferase activity, with improved
 PT productivity.

PS Claim 7; Page 20-21; 26pp; Japanese.

XX

This Page Blank (uspto)